

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2005, 18:54:32 ; Search time 4770 Seconds
(without alignment)
10613.308 Million cell updates/sec

Title: US-10-630-636-1
Perfect score: 1330
Sequence: 1 gtgcgagatcaccaccgca.....gacccaaaaaaaaaaaaa 1330

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hc1:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_g881:*
9: gb_g882:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	447	33.6	478	7	CF326660 JMT1--06-
2	432.2	32.5	461	9	CL794965 OR CBA000
3	370.2	27.8	689	5	BQ483926 WHE514_C
4	352.6	26.5	359	9	CL728595 OR BBA005
5	311.2	23.4	815	6	CA185031 SCBS1309
6	303.4	22.8	618	9	CL782098 OR BBA009
7	274.4	20.6	884	9	CG257740 OGXC048TH
8	254.6	19.9	689	9	CG710311 OGDH25TC
9	252.2	19.7	804	9	CG366978 OGAM96TV
10	252.2	19.7	853	9	CG362076 OGICL41TH
11	252.2	19.7	933	9	CG219911 OG1AD74TH
12	250.6	19.6	807	9	CL968676 OaIFCC017
13	254.8	19.2	740	9	CG751932 ZMWB013
14	253.6	19.1	468	4	BM038003 U001D12 O
15	251.2	18.9	653	9	CG752461 ZMWB013
16	242.4	18.2	560	9	CL620317 OR BBA001
17	240.2	18.1	921	9	CG257748 OGXC048TV
18	233.8	17.6	759	8	BZ525067 OGAF180TC
19	224.4	16.9	606	6	CA235113 SCCEFLA09
20	223	16.8	874	9	CG219919 OG1AD74TH
21	223	16.8	912	9	CG366964 OGAM96TV
22	220.6	16.6	914	9	CG362086 OGICL41TV
23	218.2	16.4	515	9	CL707638 OR BBA002
24	212.8	16.0	763	9	CG693661 OGFAU36TV

25	204.4	15.4	421	8	BH874076 hp48c11.9
26	204.4	15.4	620	6	CA273712 SCRLSD100
27	199	15.0	921	9	CL297228 ZMWB000
28	198.8	14.9	503	6	CA151962 SCJPR2201
29	198.2	14.9	661	4	B1954597 HVSMEM001
30	195.2	14.7	705	6	CA243768 SCQSF1407
31	191.6	14.4	454	6	CA692850 wlm96_PkO
32	190.8	14.4	643	6	CA233892 SCSCFLA0C
33	186.6	14.0	639	7	CF428085 PH1_12 GO
34	185.6	14.0	628	2	BE356185 DG1_123 F
35	185	13.9	348	4	BM500306 PACG00000
36	180.2	13.5	477	4	BM376711 EBEM05_SQ
37	180.2	13.5	666	4	BM073319 MEST64-C1
38	180.2	13.5	725	4	BM075490 MEST355-H
39	179.6	13.5	900	9	CL960237 OaIFCC004
40	178.6	13.4	601	4	BM333529 MEST157-A
41	175.8	13.2	613	6	CA292935 SCRLFL002
42	175.2	13.1	517	9	CG710318 OGDH26TM
43	174.6	13.1	741	7	CO533588 3530_1_22
44	172.2	12.9	586	7	CN010050 WHE3865 H
45	171.8	12.9	690	7	CO524463 3530_1_16

ALIGNMENTS

RESULT 1
CF326660
LOCUS
DEFINITION
JMT1--06-K20.g1 AtUMT-overexpressing transgenic rice lambda phage
CDNA library (JMT1) Oryza sativa (japonica cultivar-group) cDNA
clone JMT1--06-K20, mRNA sequence.
CF326660
ACCESSION
CF326660.1 GI:33801577
VERSION
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eriactroideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 478)
Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
Large-scale Sequencing Analysis of Rice ESTs
Unpublished (2003)
COMMENT
Contact: Nahm B.H.
Genomics and Genetics Institute, Greengene Biotech Inc., Division
of Bioscience and Bioinformatics, Wonyngi University
Yongin, Kyeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.
Location/Qualifiers
1..478
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:3947"
/clone="JMT1--06-K20"
/tissue_type="leaf"
/dev_stage="14 days after germination"
/lab_host="E. coli SOLR"
/clone_lib="AtUMT-overexpressing transgenic rice lambda
phage CDNA library (JMT1)"
/note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
XhoI; CDNA was inserted into lambda Uni-ZAP XR vector at 5'
end with EcoRI and 3' end with XhoI site. mRNA was
prepared from Arabidopsis jasmonate Carboxyl
methyltransferase overexpression line."

ORIGIN
Query Match 33.6%; Score 447; DB 7; Length 478;
Best Local Similarity 98.3%; Pred. No. 4.2e-87;

```
Matches 462; Conservative 0; Mismatches 6; Indels 2; Gaps 1;
QY 475 CTTCGCGGGCTGACAAAGTTTCGCAAGGGGAGACTGCGGAGCATCTCGGCACTTCGT 534
    |||
Db 9 CTTCGCGGGCTGACAAAGTTTCGCAAGGGGAGACTGCGGAGCATCTCGGCACTTCGT 68
QY 535 CATCTCGCGGACCGCAACGAGGTGGAGACCAAGCGGCACAAGTACTTATTCGCGCTCA 594
    |||
Db 69 CATCTCGCGGACCGCAACGAGGTGGAGACCAAGCGGCACAAGTACTTATTCGCGCTCA 128
QY 595 CTTCATGAACCGGACCGCGCGCTCCAGCATTCACGATCAGCAGCATCAGCGCTCACCGCGG 654
    |||
Db 129 CTTCATGAACCGGACCGCGCGCTCCAGCATTCACGATCAGCAGCATCAGCGCTCACCGCGG 188
QY 655 CGATCAGGTGCGCGCGAGAGAGGCGCGCGCATCAGCGGCACACAGGCCACGGGCAACCC 714
    |||
Db 189 CGATCAGGTGCGCGCGAGAGAGGCGCGCGCATCAGCGGCACACAGGCCACGGGCAACCC 248
QY 715 CGGCGG--CGGCGGCGCTGGGCGCGCGCGCATGAGACCAACCAACCAACCGCGGCG 772
    |||
Db 249 CGGCGGCGGAGAGAGCTGGGCGCGCGCGCATGAGACCAACCAACCAACCGCGGCG 308
QY 773 GGGCGCGCGCGCGCGCATGAGACGCGCGCGCGCATGAGGAGCAACCGCGTCCGCGGCG 832
    |||
Db 309 GGGCGCGCGCGCGCGCATGAGACGCGCGCGCGCATGAGGAGCAACCGCGTCCGCGGCG 368
QY 833 CACATGATGCGCGCGCGCGCTCGGAGCGCGCGGTGTTCGCGCGGCGCAACCGCGCTAC 892
    |||
Db 369 CACATGATGCGCGCGCGCGCTCGGAGCGCGCGGTGTTCGCGCGGCGCAACCGCGCTAC 428
QY 893 GTGTGTGCGCGCTCGGCTACCGCGGCGCTCGGCGCAAGTGCACCAATGACG 942
    |||
Db 429 GTGTGTGCGCGCTCGGCTACCGCGGCGCAACCGCGCAAGTGCACCAATGACG 478

RESULT 2
CT794965/c 461 bp DNA linear GSS 06-AUG-2004
LOCUS OR_CBA00040J23.f OR_CBA Oryza rufipogon genomic clone OR_CBA00040J23
DEFINITION 5', genomic survey sequence.
ACCESSION CT794965
VERSION CT794965.1 GI:51016981
KEYWORDS GSS.
SOURCE Oryza rufipogon
ORGANISM Oryza rufipogon
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 461)
Kim,H., Yu,Y., Wisesotaki,M., Yost,D., Stum,D., Rao,K., Luo,M.,
Jecty,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and
Wang,R.
OMAP project
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Plate: 0004 row: J column: 23
Seq primer: TAA TAC GAC TCA CTA TAG GG
Class: BAC ends.
Location/Qualifiers
1. 461
/organism="Oryza rufipogon"
/mol_type="genomic DNA"
/db_xref="taxon:4529"
/clone="OR_CBA00040J23"
```

```
ORIGIN
Query Match 32.5%; Score 432.2; DB 9; Length 461;
Best Local Similarity 98.9%; Pred. No. 7.1e-84;
Matches 456; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
/cisue_type="young leaves"
/dev_stage="2 week old seedlings"
/lab_host="PH10B T1 phage resistant"
/clone_lib="OR_CBA"
/notes="Vector: pGIBAC1; Site 1: HindIII; Site 2: HindIII;
drk treated 36 hrs before harvest"

QY 720 CGGCGGCGCTGGGCGCGCGGCGCA--TGAAGCAACCAACCA--CCACACCGCGGCGGCGG 777
    |||
Db 461 CGGCGGCGCTGGGCGCGCGGCGCATTAAGACACACACACACACACACACCGCGGCGGCGG 402
QY 778 GCGCGGCGCATGCGCATGACAGCGCGCGGCGCATGAGGCGCAACCGCGTCCGCGGCGACAT 837
    |||
Db 401 GCGCGGCGCATGCGCATGACAGCGCGCGGCGCATGAGGCGCAACCGCGTCCGCGGCGACAT 342
QY 838 GGTGCGCGCGCGCGTGGAGCGCGGCGGTGTTCCCGCGGCGACGCGCGTACGTCGT 897
    |||
Db 341 GGTGCGCGCGCGCGTGGAGCGCGGCGGTGTTCCCGCGGCGACGCGCGTACGTCGT 282
QY 898 GCGCGTGGCTACCGCGGCGCTTCGCGGCGCAAGATGACCAATGACGCGCGCATGAGACGAGCA 957
    |||
Db 281 GCGCGTGGCTACCGCGGCGCAACCGCGCAAGATGACCAATGACGCGCGCATGAGACGAGCA 222
QY 958 TGAGCAGCATTTCTTCCTCCTCTTCTTGATGTCAATCTTGATTTGTTTGTGTAGT 1017
    |||
Db 221 TGAGCAGCATTTCTTCCTCCTCTTCTTGATGTCAATCTTGATTTGTTTGTGTAGT 162
QY 1018 CGCGGCGTACGTCGCTGATCATCTTGTCTTCTGACAACTCTACTAATGTAACATAC 1077
    |||
Db 161 CGCGGCGTACGTCGCTGATCATCTTGTCTTCTGACAACTCTACTAATGTAACATAC 102
QY 1078 ATAGATCAGATGCGCAGAGTGCAGGAGTGGGATTAAGCGCAATGTAAGTATTTT 1137
    |||
Db 101 ATAGATCAGATGCGCAGAGTGCAGGAGTGGGATTAAGCGCAATGTAAGTATTTT 42
QY 1138 GCTGACTGTTTTCGATGATCATCATGTCACCGCGTGAAA 1178
    |||
Db 41 GCTGACTGTTTTCGATGATCATCATGTCACCGCGTGAAA 1

RESULT 3
BQ483926 689 bp mRNA linear EST 03-JUN-2002
LOCUS WHE3514.C07_E14ZS wheat unstressed root cDNA library Triticum
DEFINITION WHE3514.C07_E14ZS wheat unstressed root cDNA library Triticum
ACCESSION BQ483926
VERSION BQ483926.1 GI:21319862
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 689)
Anderson,O.D., Chao,S., Close,T.J., Crossman,C., Fenton,R.D.,
Lazo,G.R., Nguyen,H.T., Rauech,C.U., Wilson,C., Woo,J. and Zhang,D.
The structure and function of the expressed portion of the wheat
genomes - Unstressed root cDNA library
Unpublished (2002)
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
```


source
1. .618
/organism="Oryza rufipogon"
/mol_type="genomic DNA"
/db_xref="taxon:4529"
/clone="OR_B86038B05"
/issue_type="Young leaves"
/lab_host="DH10B-T1 phage resistant"
/clone_lib="OR_B86"
/note="Vector: pGIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match 22.8%; Score 303.4; DB 9; Length 618;
Best Local Similarity 97.8%; Pred. No. 1e-55;
Matches 318; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy 1001 TTTGTTCTTTGTGATGTC -CCGGCTCATGCTCCCTGATCATCTTTGTTCTTCTCAATC 1059
Db 610 TTTGTTCTTTGTGATGTCGCCCGCTCATGCTCCCTGATCATCTTTGTTCTTCTCAATC 551

Qy 1060 TCACTAATGTAACATACATAGATCAGATGCCAAGAGCGAGGATTGGGATTTAAAGGC 1119
Db 550 TCACTAATGTAACATACATAGATCAGATGCCAAGAGCGAGGATTGGGATTTAAAGGC 491

Qy 1120 GAATAGTAAGTAATTTGCTGACTGTTTGCAGTATCATCAGTACACCCGGTGAAG 1179
Db 490 GAATAGTAAGTAATTTGCTGACTGTTTGCAGTATCATCAGTACACCCGGTGAAG 431

Qy 1180 CTTAGCTCCAAATGTGATGTAATTAAGACGGGCTTCCGTAAGTGTTGGCGCGCATGA 1239
Db 430 CTTAGCTCCAAATGTGATGTAATTAAGACGGGCTTCCGTAAGTGTTGGCGCGCATGA 371

Qy 1240 TGATCTTGACAGGGGTGCAATTAAGGATTTGATTTCCATTTTGTGATGTAATTTGGCAA 1239
Db 370 TGATCTTGACAGGGGTGCAATTAAGGATTTGATTTCCATTTTGTGATGTAATTTGGCTTA 311

Qy 1300 CTGCTCATTTGACCAAAAAAAAA 1324
Db 310 CTGCTCATTTGACCAAAAAAAAACATATA 286

RESULT 7
CG257740 884 bp DNA linear GSS 25-AUG-2003
LOCUS OGC048TH ZM 0.7.1.5_KB zea mays genomic clone ZMMBMA0650G23,
DEFINITION genomic survey sequence.
ACCESSION CG257740
VERSION CG257740.1 GI:34164524
KEYWORDS GSS.
SOURCE zea mays
ORGANISM zea mays
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 884)
Whiteley, C.A., Quackenbush, J., Van Aken, S., Utecherack, T.,
Renwick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OGC048TV
Contact: Cathy Whiteley
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteley@tigr.org
Seq primer: TR
Class: sheared ends.
location/Qualifiers
1. .884
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="873"

source
/db_xref="taxon:4577"
/clone="ZMMBMA0650G23"
/clone_lib="ZM 0.7.1.5_KB"
/note="Vector: pBCK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 20.6%; Score 274.4; DB 9; Length 884;
Best Local Similarity 71.4%; Pred. No. 2.3e-49;
Matches 528; Conservative 0; Mismatches 96; Indels 116; Gaps 8;

Qy 290 GGGAGGAGTCCCGGCGCCCGCAGAGACCCGAGCCCGCGCTCCAGAGAC 349
Db 869 GAGAGTGTCTCCGCGAGCGCCGAGGGCTCCGATCCGCGCTCCGCCCAAGGAC 810

Qy 350 GGCGGA-----CACGCGCCCAAGAGCCCAAGGCGGCGCGC 388
Db 809 GCTGAGAGAGAGAGAGAGGCGGTGGGCACGAGCGCAAGAACGAGGCGGTCCGC 750

Qy 389 GGTACGACGCGCGCAAGAGCTGCTCCAGGCGGAGCGAGAGAGCGCAAGGCGATCCA 448
Db 749 GTT---GACGTGGGAAAGTAGCTTAAGCTAGACAGAGCGCGAGAGGATCCG 693

Qy 449 TGAAGAGAGAGACAGGC----- 470
Db 692 TGAAGAGAGAGACAGGTTCTAGAGATCTCCACGTTCTCGCTTCAAAAGCT 633

Qy 471 -----TGTTCTGCTGGGCTGGAACAAGTTCCGCAAGGCG 506
Db 632 TACACCTTTTGTGGGTGAGGTGAGTTGTTCTTGCTGGGCTTGAACAAGTTCCGCAAGGCG 573

Qy 507 ACTGGCGAGCATCTCGCGCAACTTGTCATCTCGCGAGCGCAAGCGAGGTGGCGAGC 566
Db 572 ACTGGCGAGCATCTCGCGCAACTTGTCATCTCGAGCGCGCAAGCGAGGTGGCGAGC 513

Qy 567 ACGGCAAGAGTACTTCTGCTGCTCAATCCATGAACCGGACCGCGCGCTTCAAGCA 626
Db 512 ACGGCAAGAAATACTTCTGCTGCTCAATCCATGAACCGGACCGCGCGCTTCAAGCA 453

Qy 627 TCCACGATCAACAGGCTCACCGCGCGATCAAGTCCGCGCAAGGCGCGCCCGA 686
Db 452 TCCACGATCAACAGGCTGAGCGCGCGCAAGTCCGCGCGC-----CGGCCACCA 399

Qy 687 TCACGCGCACCAAGCGCAACCCCGCGCGCGCGCTGGGCGCGCGCGGCAATGA 746
Db 398 TCACGCGCGCGCGCAACCG-----CGGAGCGATGCCATGATCCCGGCAATGA 345

Qy 747 AGCACACACACACACACCGCGCGCGCGCGCGCGCGCGCATGTACAGCGCG 806
Db 344 AGCACACACACACACAC-----CGGCTCGCGATGGGATGTACCGGCGACG 297

Qy 807 CGCCCATGGGCGCACCGCGCGCGCGCGCATAGT---GCCCGCGCGCTCGCACCGCG 863
Db 296 CGCCCATGGGCGCACCGCGCGCGCGCGCATAGTGGGCGCGCGCGCGCTCGCACCGCG 237

Qy 864 TGGTGTCCCGCGCGG---CCAGCGCGCGTACGTCGTCGCGCGCGTACCGCGCGCTC 920
Db 236 TCATGTTCCCGCGCGGTCATCTCGGCTTAGTGTGACAGTGGATACCGCGGCGCGC 177

Qy 921 CGGCAAGATGACCAATGA 940
Db 176 CGGCAAGATGACCAATGA 157

RESULT 8
CC710311/c 689 bp DNA linear GSS 19-JUN-2003
LOCUS OGDH26TC ZM 0.7.1.5_KB zea mays genomic clone ZMMBMA0145F04,
DEFINITION genomic survey sequence.
ACCESSION CC710311
VERSION CC710311.1 GI:32115087
KEYWORDS GSS.
SOURCE zea mays

ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 689)
REFERENCE Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
AUTHORS Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.M., Nunez, A., Robbins, D. and Lakey, N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other GSSs: OGDH26TM
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
source
1. .689
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZM0.7.1.5_KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 19.9%; Score 264.6; DB 9; Length 689;
Best Local Similarity 71.9%; Pred. No. 3.1e-47;
Matches 429; Conservative 0; Mismatches 134; Indels 34; Gaps 5;
466 CAGGCTGTTCTGCTGGGCTGGAACAAGTTGCGCAAGGGGACTGGCGAGCATCTGCG 525
466 |||||
576 CAGGTTTCTTGTGTGGCTTGGACAAAGTTGCGCAAGGGGCACTGGCGAGCATCTGCG 517
576 |||||
526 CAACTTGATCTCGCGGAGCGCAAGTGGCGAGCGACGCGAGAGTAATTCTAT 585
526 |||||
516 CAATTGATCTCGCGGAGCGCGAGTGGCGAGCGAGCAAGAAATCTTCAT 457
516 |||||
586 CCGCTCACTCAATGAACCGGCGCGCGCTGCTCAAGATCCAGCATCACCAGCT 645
586 |||||
456 CCGCTCACTCAATGAACCGGCGCGCGCTGCTCAAGATCCAGCATCACCAGCT 397
456 |||||
646 CACGCGCGGATGAGTGGCGCGGAGCGAGGCGCGCGCATCAAGCGACAGGCGAC 705
646 |||||
396 GACGCGCGGAGTGGCGCGGAGCGCGCGCGCATCAAGCGCGCGCGCGCGCG 337
396 |||||
706 GAGGCAACCCGCGCGCGCGCTGAGCGCGCGCGGAGTGAAGCAACCAACCA 765
706 |||||
336 GGG-----GGCATGCGATGGGGCGGGCGGCGATGAAGCAACCAACCC----- 293
336 |||||
766 CCGCGGCGGCG 825
766 |||||
292 -----CGGTTCCGCGCGATGGGCGATGTAAGGCGCGCGCGCGCGCGCG 244
292 |||||
826 CGCGCGCGCATGGT---GCCCGCGCGCGTGGCGAGCGCGGTTGTTCCCGCGCGCG 880
826 |||||
243 CGCGCGCGCATGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 184
243 |||||
881 -CAGCGCGCGTACGTCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 939
881 |||||
183 TCACTGCGCGTATCTCTGCGCGGATGACCGCGCGCGCGCGCGCGCGCGCGCG 124
183 |||||
940 ACGCGCGATGACGG-AACATGACGATTTCTTCTCTCTCTCTCTCTCTCTCTCT 998
940 |||||
123 ATCTCTACGACCGCGCGATTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 64
123 |||||
999 GATTGTCTTGTGTAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1055
999 |||||

Db 63 GATTGTCTTGTGTAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7
RESULT 9
CG366978/c 804 bp DNA linear GSS 26-AUG-2003
LOCUS OGMAM96TV ZM 0.7.1.5_KB Zea mays genomic clone ZMBMA0520023,
DEFINITION genomic survey sequence.
ACCESSION CG366978
VERSION CG366978.1 GI.34284245
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 804)
REFERENCE Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
AUTHORS Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.M., Nunez, A., Robbins, D. and Lakey, N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other GSSs: OGMAM96TV
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
source
1. .804
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZM0.7.1.5_KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 19.7%; Score 262.2; DB 9; Length 804;
Best Local Similarity 74.9%; Pred. No. 1.1e-46;
Matches 408; Conservative 0; Mismatches 103; Indels 34; Gaps 5;
466 CAGGCTGTTCTGCTGGGCTGGAACAAGTTGCGCAAGGGGACTGGCGAGCATCTGCG 525
466 |||||
689 CAGGTTTCTTGTGTGGCTTGGACAAAGTTGCGCAAGGGGCACTGGCGAGCATCTGCG 630
689 |||||
526 CAACTTGATCTCGCGGAGCGCAAGTGGCGAGCGACGCGAGAGTAATTCTAT 585
526 |||||
629 CAATTGATCTCGCGGAGCGCGAGTGGCGAGCGAGCAAGAAATCTTCAT 570
629 |||||
586 CCGCTCACTCAATGAACCGGCGCGCGCGCTGCTCAAGATCCAGCATCACCAGCT 645
586 |||||
569 CCGCTCACTCAATGAACCGGCGCGCGCGCTGCTCAAGATCCAGCATCACCAGCT 510
569 |||||
646 CACGCGCGGATGAGTGGCGCGGAGCGAGGCGCGCGCATCAAGCGCGACAGGCGAC 705
646 |||||
509 GACGCGCGGAGTGGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 450
509 |||||
706 GAGGCAACCCGCGCGCGCGCTGAGCGCGCGCGGAGTGAAGCAACCAACCAACCA 765
706 |||||
449 GGG-----GGCATGCGATGGGGCGGGCGGCGGATGAAGCAACCAACCAAC 406
449 |||||
766 CCGCGGCGGCG 825
766 |||||
405 -----CGGTTCCGCGCGATGGGCGATGTAAGGCGCGCGCGCGCGCGCGCG 357
405 |||||
826 CGCGCGCGCATGGT---GCCCGCGCGCGTGGCGAGCGCGGTTGTTCCCGCGCGCG 880
826 |||||

Db 356 CGCCGGGACATGGTGGCGCGCGCGCGCGCATGTTCCCGCCGGGCA 297
Qy 881 -CAGCGCGGTACGTGTCGCGCTACCGCGCGCTCGCGGCAAGATGACCAATG 939
Db 296 TCATCGCGCTTATGTCCTCGCGGTGGATACCGCGCGCATGACCAAGATGACCAATG 237
Qy 940 ACGCGCCATGACCG-ACATGAGCAGATTTCTTCTCTCTTCTTATGTCATCTT 998
Db 236 ATCGCTACGACGCGCATTTGCGGAGTGGCGAGCTGATCGCTTGTCTCTTCTT 177
Qy 999 GATTT 1003
Db 176 GATTT 172

RESULT 10
CG362076 853 bp DNA linear GSS 26-AUG-2003
LOCUS OG1CL41TH ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0730H10,
DEFINITION genomic survey sequence.
ACCESSION CG362076
VERSION CG362076.1 GI:34279343
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 853)
WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Reinick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Numberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OG1CL41TV
Contact: Cathy WhiteLaw
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteLaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source location/Qualifiers
1..853
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0730H10"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 19.7%; Score 262.2; DB 9; Length 853;
Best Local Similarity 74.9%; Pred. NO. 1.1e-46;
Matches 408; Conservative 0; Mismatches 103; Indels 34; Gaps 5;

Qy 466 CAGGCTGTTCTTGTGGGCTGGAACAAGTTGGCAAGGGGAGCTGCGGAGCATCTCGCG 525
Db 264 CAGGTTGTTCTTGTGGGCTGGAACAAGTTGGCAAGGGGAGCTGCGGAGCATCTCGCG 323
Qy 526 CAATTTGCTATCTTCGGGAGCCCAAGCGAGTGGCGGACGCGGAGAACTTCTAT 585
Db 324 CAATTTGCTATCTTCGGGAGCCGAGCGAGTGGCGGACGCGGAGAACTTCTAT 383
Qy 586 CCGGCTCAATCTTCATGACGCGGAGCGCGGCTCGCATGACGATCAGCAGCGT 645
Db 384 CCGGCTCAATCTTCATGACGCGGAGCGCGGCTCGCATGACGATCAGCAGCGT 443
Qy 646 CACGCGCGGATGATGATCGCGCGGAGCAGGCGCGCGCATCAGGCGCAC 705

Db 444 GACCGCGGACAGGTGGCCAGCGCGCGCGCGCCCATCAAGGAGCGCGCGCGCGC 503
Qy 706 GGGCAACCCCGCGGCGGCGGCTGAGGCGCGCGCGGATGAGGACCAACCAACCA 765
Db 504 GGG-----GGGATGCGGATGAGGCGCGGCGGATGAGGACCAACCAACCA 547
Qy 766 CCGGCGGCGCGCGCGCGCGCCATGATGATGAGGCGCGCGCGCGCGCGCGCGT 825
Db 548 -----CGGTTCCGCGGATGAGGCGGATGAGGCGCGCGCGCGCGCGCGCGT 596
Qy 826 CCGCGGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 880
Db 597 CGCGCGGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 656
Qy 881 -CAGCGCGGTACGTGTCGCGCTACCGCGCGCTCGCGGCAAGATGACCAATG 939
Db 657 TCATCGCGCTTATGTCCTCGCGGTGGATACCGCGCGCATGACCAAGATGACCAATG 716
Qy 940 ACGCGCCATGACCG-ACATGAGCAGATTTCTTCTCTCTTCTTATGATGTCATCTT 998
Db 717 ATCGCTACGACGCGCATTTGCGGAGTGGCGAGCTGCGGAGCTGATCGCTTGTCTTCTT 776
Qy 999 GATTT 1003
Db 777 GATTT 781

RESULT 11
CG219911/c 933 bp DNA linear GSS 22-AUG-2003
LOCUS OG1AD74TH ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0715N04,
DEFINITION genomic survey sequence.
ACCESSION CG219911
VERSION CG219911.1 GI:34119799
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 933)
WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Reinick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Numberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OG1AD74TV
Contact: Cathy WhiteLaw
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteLaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source location/Qualifiers
1..933
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0715N04"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 19.7%; Score 262.2; DB 9; Length 933;
Best Local Similarity 74.9%; Pred. NO. 1.1e-46;
Matches 408; Conservative 0; Mismatches 103; Indels 34; Gaps 5;

Qy 466 CAGGCTGTTCTTGTGGGCTGGAACAAGTTGGCAAGGGGAGCTGCGGAGCATCTCGCG 525

Db 786 CAGGTTCTTGTGGGTGTGACAAAGTTCCGGAGAGGCGGAGCTGGCGGACATCTCGCG 727

Qy 526 CAACCTTGTATCTCGCGGACGCCCAACGACGAGTGGCGAGCCACGCGCAGAATCTTCAT 585

Db 726 CAACCTTGTATCTCGCGGACGCCCAACGACGAGTGGCGAGCCACGCGCAGAATCTTCAT 667

Qy 586 CCCCCTCAATCTCATGAAACCGGACCGCGCCCGCTCCAGAGATCCACGACATCCACCGCT 645

Db 666 CCGCTCAATCTCATGAAACCGGACCGCGCCCGCTCCAGAGATCCACGACATCCACCGCT 607

Qy 646 CACCGCGCGGATCAGGTGCGCGGACGAGAGGCGCCCGATCAACGCGCAGCGCAC 705

Db 606 GACCGCGCGGAGAGTGGCCACGCGCGCGCGCCCATACGAGGCGCGCGCCGACCGCGCGC 547

Qy 706 GGGCAACCCCGCGCGCGCGCTGGGCGCCCGCGGATGAAGACCAACCAACCA 765

Db 546 GGG-----GGCGATGCGGATGGGGCGGGGCGGATGAAGACCAACCAACCA----- 503

Qy 766 CCGGCGCGGCGCGCGCGCGCTCCATGCCCATGTACAGCGCGCGCCCATGGGCGACCCCGT 825

Db 502 -----CGGTTCCGCGCGAGTGGGCGATGTACGAGCGACGCGCCCATGGGCGACCCCGT 454

Qy 826 CGCGCGCGCATGATGAT---GCCCGCGCGCGTGGGCGACGCGCGTGGTGTTCGCGCGGCG-- 880

Db 453 CGCGCGCGCATGATGATGCGCGCGCGCGCGCGCGCTGCGACGCGCGCATGTTCCCGCGGCGCA 394

Qy 881 -CAGCGCGCGTACGTGTCGCGCGCTGCGGCTACCGCGCGCGCTCCGCGCAAGATGACCAAG 939

Db 393 TCATCGCGCTATGTCTGCGCGGATGACCGCGCGCGCGCGCAAGATGACCAAGTGCACGATG 334

Qy 940 ACGCGCGCATGACGG-ACATGAGAGCATTTCTTCCTCCCTTCTTGTGATGTCAACTT 998

Db 333 ATGCTACGACCGCGCATTTGCGAGCTGCGCGAGCTGATCCCTTGTCTCTTCTTCTT 274

Qy 999 GATTT 1003

Db 273 GATTT 269

RESULT 12
CL968676 807 bp DNA linear GSS 21-SEP-2004
LOCUS OeIFCC017023 Oryza sativa Expressed Library Oryza sativa (indica
DEFINITION cultivar-group) genomic, genomic survey sequence.
ACCESSION CL968676
VERSION CL968676.1 GI:52391986
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eriarthroideae; Oryzaceae; Oryza.
AUTHORS 1 (bases 1 to 807)
Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G.K.S., Deng, X.W. and Wang, J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
Unpublished (2004)
JOURNAL Contact: Chen Chen
COMMENT Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-8048676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.

FEATURES
source
1..807
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"

ORIGIN
Query Match 19.6%; Score 260.6; DB 9; Length 807;
Best Local Similarity 83.9%; Pred. No. 2, 4e-46;
Matches 323; Conservative 0; Mismatches 49; Indels 13; Gaps 2;
/clone_11b="Oryza sativa Expressed Library"
/note="Oryza sativa exon trapped genomic sequences"

Qy 362 CGGACGACGCGCAAGGCGCGCGCGGATGACGCGCGGCAAGAGCTGCTCCAGGCG 421

Db 436 CGTGGCGGAAGACCGGATGTCTGCGGAGTACGTGCTGCGCAAGATCTGCTCCAGTGC 495

Qy 422 GAGCAGAGAGGCGGCGCAAGGCGCATCCATGAGCGAGGAAGAGACAGGCTGCTTGGCTG 481

Db 496 GAGCAGAGAGGCGGCGCAAGGCGCATCCATGAGTGGAGAGAGAGCAGAGGCTGTTGCTG 555

Qy 482 GGGCTGACAAAGTTCGCGCAAGGCGGAGTGGCGGAGATCTGCGCAACTTCTGATCTCG 541

Db 556 GGGCTGAGCAAGTTCGCGCAAGGCGGAGTGGCGGAGATCTGCGCAACTTCTGATCTCG 615

Qy 542 CGGACGCGCAAGGCGGAGTGGCGGAGCGGCGGAGAGTACTTATCTGCGCTCAACTCGATG 601

Db 616 CGGACGCGCGGCGGAGTGGCGGAGCGGCGGAGAGTACTTATCTGCGCTCAACTCGATG 675

Qy 602 AACCGGACCGCGCGCGCTCCAGCATCCAGCATCAACGAGTCAACCGCGGAGTCAAG 661

Db 676 AACCGGATGTGTGCGCTCCAGCATCCAGCATCAACGAGTCAACCGCGGAGTCAAG 732

Qy 662 GTGCGCGCGGAGGCGCGCGCGGATCACCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 721

Db 733 GTTGGCGCGGAGAGGCG-----CCGATCACCGGCGGAGGCGGCGGCGGCGGCGGCGG 782

Qy 722 GCGGCGCTGGGCGCGCGCGGCGATGA 746

Db 783 GCGGCGCTGGGCGCGGCGGCGGCGATGA 807

RESULT 13
CC751932 740 bp DNA linear GSS 25-JUN-2003
LOCUS ZMMBB0133C24.r ZMMBB Zea mays genomic clone ZMMBB0133C24 3',
DEFINITION genomic survey sequence.
ACCESSION CC751932
VERSION CC751932.1 GI:32211515
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
AUTHORS 1 (bases 1 to 740)
Yu, Y., Kim, H.R., Hatfield, J., Soderlund, C., Bharti, A.K., Messing, J.
and Wing, R.
Sequencing of the maize genome
Unpublished (2003)
JOURNAL Contact: Rod Wing
COMMENT Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: T7
BACKWARD: M13r
Plate: 0133 row: C column: 24
Seq primer: M13r
Class: BAC ends.

FEATURES
source
1..740
/organism="Zea mays"
/mol_type="genomic DNA"

/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZM8B013324"
/lab_host="DH10B"
/clone_lib="ZM8Bb"
/note="Vector: pBelOBAC11; Site 1: HindIII; Site 2:
HindIII; Zea mays L. ssp. mays"

ORIGIN

Query Match 19.1%; Score 254.8; DB 9; Length 740;
Best Local Similarity 73.9%; Pred. No. 4.4e-45;
Matches 403; Conservative 0; Mismatches 108; Indels 34; Gaps 5;

Qy 466 CAGGCTGTTCTTGTCTGCGCTGAGACAGTTGCGCAAGGCGGACTGCGGAGCATCTCCG 525
Db 13 CAGGTTGTTCTTGTTCCTTTGGACAGTTGCGNAAGGCGACTGCGGAGCATCTCCG 72
Qy 526 CAATTCGTCTATCTTGGCGAGCGCAAGCGAGGTGGCGACCGCGCAAGAACTTCAT 585
Db 73 CAATTCGTCTATCTTGGCGAGCGCGACCGAGGTGGCGACCGCGCAAGAACTTCAT 132
Qy 586 CGGCTCACTCCATGAGACCGGACCGCGCTCCAGCATCCAGCATCAGACAGGT 645
Db 133 CGGCTCACTCCATGAGACCGGACCGCGCTCCAGCATCAGCATCAGACAGGT 192
Qy 646 CACGCGCGGATCAGTCTGCGCGCAGCAGGCGCGCGATCAGCGCGCAAGCGCAC 705
Db 193 GACCGCGCGAGTGGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 252
Qy 706 GGGCAACCCCG 765
Db 253 GGG-----GGGATGCGGATGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 296
Qy 766 CCG 825
Db 297 -----CGGTTCCGCGGATGGCGATGACGCGCGCGCGCGCGCGCGCGCGCG 345
Qy 826 CGCGCGCGCATAGT---GCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 880
Db 346 CGCGCGCGCATAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 405
Qy 881 -CAGCGCGCGTATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 939
Db 406 TCATCGCGCGTATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 465
Qy 940 ACGGCGCATGAGCG-AATGAGCAGCATTTCTTCTCTCTTCTTCTTCTTCTTCTT 998
Db 466 ATGCTTACGACCGCATTTGCGGAGCTGCGCGCGCGCGCGCGCGCGCGCGCG 525
Qy 999 GATT 1003
Db 526 GATT 530

RESULT 14 468 bp mRNA linear EST 06-NOV-2001
BM038003
LOCUS U001D12 Oryza sativa mature leaf library induced by M.grisea Oryza
DEFINITION sativa cDNA clone U001D12, mRNA sequence.
ACCESSION BM038003
VERSION BM038003.1 GI:16753624
KEYWORDS EST.
SOURCE Oryza sativa
ORGANISM Oryza sativa
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Bhartholidae; Oryzaceae; Oryza.
1 (bases 1 to 468)
AUTHORS Dong H.T., Li D.B., Zhang X.F., Dai C.G., Sun L.X., Pei Y.X.,
Wu H.F., Jiang Y.X., Yu F.C., Gao Q.K. and Lou Y.C.
TITLE A Gene Expression Screen in Oryza sativa
JOURNAL Unpublished (2001)
COMMENT Contact: Haitao Dong, Dehao Li

Bioinformatics and Gene Network Research Group
Zhejiang University
Kaixuan Road 268#, Hangzhou, Zhejiang, P.R. China
Tel: 0086-571-86892051
Fax: 0086-571-868961525
Email: webmaster@estarray.org, URL: http://www.estarray.org
Seq primer: M13 forward primer.
Location/Qualifiers

FEATURES

SOURCE

1. 468
/organism="Oryza sativa"
/mol_type="mRNA"
/db_xref="taxon:4530"
/clone="U001D12"
/issue_type="leaf"
/dev stage="Mature stage"
/clone_lib="Oryza sativa mature leaf library induced by
M.grisea"
/note="Vector: pSport2"

ORIGIN

Query Match 19.1%; Score 253.6; DB 4; Length 468;
Best Local Similarity 92.6%; Pred. No. 7.8e-45;
Matches 311; Conservative 0; Mismatches 19; Indels 6; Gaps 4;

Qy 993 AATCTGATTTGTTCTTGT--GTAGTCGCGCGCTCATCGTCCGATCATCTTGTCTT 1050
Db 7 AATCTGATTTGTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 66
Qy 1051 CTCACATCTCATATGTAATGTAACATATCATATGATGATGATGATGATGATGATG 1108
Db 67 CTCACATCTCATATGTAATGTAACATATCATATGATGATGATGATGATGATGATG 126
Qy 1109 GGATTAAGCGGATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 1168
Db 127 GGATTAAGCGGATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 186
Qy 1169 CCGGTAAGCTTATGCTCCAAATGATGATGTAATGTAATGTAATGTAATGTAATG 1228
Db 187 CCGGTAAGCTTATGCTCCAAATGATGATGTAATGTAATGTAATGTAATGTAATG 245
Qy 1229 GCGCGATGATGATCTTTCAGGCGGTGCAATTAGGATGATGATGATGATGATG 1288
Db 246 GC-CGATGATGATGATCTTTCAGGCGGTGCAATTAGGATGATGATGATGATG 304
Qy 1289 AAATTTGCCAATGTCATTTGACCAAAAAAAA 1324
Db 305 AAATTTGCTTACTGTCATTTGACCAAAACATATA 340

RESULT 15 653 bp DNA linear GSS 25-JUN-2003
CC752461
LOCUS ZM8B0133P10.f ZM8Bb Zea mays genomic clone ZM8B0133P10 5',
DEFINITION genomic survey sequence.
ACCESSION CC752461
VERSION CC752461.1 GI:32212582
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 653)
AUTHORS Yu Y., Kim H.R., Hatfield J., Soderlund C., Bharti A.K., Messing J.
and Wing R.
TITLE Sequencing of the maize genome
JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967

Fax: 520 621 9288
Email: <http://genome.arizona.edu>
PCR Primers
FORWARD: T7
BACKWARD: M13r
Plate: 0133 row: P Column: 10
Seq primer: T7
Class: BAC ends.

FEATURES
Source

Location/Qualifiers
1..653
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMMBB01.3P10"
/lab_host="DH10B"
/clone_lib="ZMMBBb"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII; Zea mays L. ssp. mays"

ORIGIN

Query Match 18.9%; Score 251.2; DB 9; Length 653;
Best Local Similarity 73.8%; Pred. No. 2,7e-44;
Matches 400; Conservative 0; Mismatches 108; Indels 34; Gaps 5;

```
QY 469 GCTGTTCTTCTGCGGCTGACAAAGTTCGGCAAGGAGGAGCTGCGAGCATTTGCGCAA 528
Db 63 GTTGTCTTCTGTTGGTCTGGACAAAGTTCGGCAAGGAGGAGCTGCGAGCATTTGCGCAA 122
QY 529 CTTTCGTATCTCGCGGACCGCAACGAGTGGCGAGCCAGCCGCGAGAACTTCATCCG 588
Db 123 CTTTCGTATCTCGCGGACCGCAACGAGTGGCGAGCCAGCCGCGAGAACTTCATCCG 182
QY 589 CCTCAACTCATGACCGGACCGCCGCTCCAGCATCCAGCATCAACAGCTCAC 648
Db 183 CCTCAACTCATGACCGGACCGCCGCTCCAGCATCCAGCATCAACAGCTCAC 242
QY 649 CGCGGCGATCAGTCCCGCGCAGCAGGCGCCCGATCAACGCGCAACAGCGCACGG 708
Db 243 CGCGGCGATCAGTCCCGCGCAGCAGGCGCCCGATCAACGCGCGCGCGCGG 302
QY 709 CAACCCCGCGCGCGCGCTGCGCGCGCGCGCGCATGAAGACCAACCAACCAACC 768
Db 303 -----GCGCATGCCGATGCGGCGCGCGCGCATGAAGACCAACCAACC----- 343
QY 769 GGGCGGCGCGCGCGCGCATGCGCATGTAACGCGCGCGCGCGCATGCGCGCACCCGTCGC 828
Db 344 -----CGGTTCCGCTCATGGGCGATGACGGGCAACCGCCCATGGGCGACCCGCTCGC 395
QY 829 CGGCGCATGAGT---GCCGCGCGCGCTCGCGCACCGCGGTGTTCCTCGCGCGGC---CA 882
Db 396 CGGCGCATGAGTGGCGCGCGCGCGCGCTCGCGCACCGCGGTGTTCCTCGCGCGGCATCA 455
QY 883 CGGCGCGTACGTCGTGCGCGCGCTACCGCGCGCTCCCGCGCAAGTGAACCAATGACG 942
Db 456 CTGCGCTATGTCCTGCGCGGTGAGTACCGCGCGCACCGCACGACGACGCGATC 515
QY 943 CGCCATGACGCG--ACATGAGACGATTTCTCTCTCTCTTTCTTGATGTCAATCTTGAT 1001
Db 516 GCCTACGACCGGCAATTCCGCGAGCTGCGCAGCTGCAATCCCTTGTCTGTTCTTTGAT 575
QY 1002 TT 1003
Db 576 TT 577
```

Search completed: July 29, 2005, 22:29:27
Job time : 4781 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2005, 15:06:02 ; Search time 785 seconds
(without alignments)
10029.625 Million cell updates/sec

Title: US-10-630-636-1

Perfect score: 1330
Sequence: 1 gtgcgagatccaccaccgca.....gacccaaaaaaaaaaaaa 1330

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1:	geneseqn1980s:*
2:	geneseqn1990s:*
3:	geneseqn2000s:*
4:	geneseqn2001as:*
5:	geneseqn2001bs:*
6:	geneseqn2002as:*
7:	geneseqn2002bs:*
8:	geneseqn2003as:*
9:	geneseqn2003bs:*
10:	geneseqn2003cs:*
11:	geneseqn2003ds:*
12:	geneseqn2004as:*
13:	geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1330	100.0	1330	12	AD000777	Ado00777 Rice Myb
2	918	69.0	918	12	AD000780	Ado00780 Rice Myb
3	135.4	10.2	1161	3	AAC34213	Aac34213 Arabidops
4	130.6	9.8	412	12	ADP93595	Adp93595 Cotton ex
5	129.2	9.7	1432	10	ADCC4636	Adc4636 Thalecres
6	129.2	9.7	1432	10	ADDS5829	Adds5829 Thalecres
7	129.2	9.7	1432	12	AD161458	Adi61458 CDNA enco
8	129.2	9.7	1432	12	AD002450	Ado02450 Thalecres
9	126.8	9.5	1119	6	ABT08115	Abt08115 Myb-relat
10	126.8	9.5	1119	12	ADH50105	Adh50105 Soybean M
11	123.2	9.3	257	12	ADQ03867	Adq03867 Maize tra
12	120.2	9.0	386	12	ADQ04396	Adq04396 Maize tra
13	119.8	9.0	315	12	ADQ03772	Adq03772 Maize tra
14	119	8.9	144	12	ADQ03861	Adq03861 Maize tra
15	118.6	8.9	867	6	ABK65275	Abk65275 Arabidops
16	118.6	8.9	867	10	ADD30621	Add30621 Plant yie
17	118.6	8.9	867	10	ADB31552	Adb31552 Plant yie
18	118.6	8.9	867	12	AD144216	Adi44216 Plant tra
19	112.8	8.5	1374	12	AD143214	Adi43214 Plant tra
20	112.8	8.5	1428	12	AD143213	Adi43213 Plant tra

21	111.8	8.4	934	3	AAC57273	Aac57273 Eucalyptu
22	111.8	8.4	1358	3	AAC56101	Aac56101 Eucalyptu
23	110.6	8.3	954	12	AD000782	Ado00782 Rice Myb
24	110.6	8.3	1287	12	AD000779	Ado00779 Rice Myb
25	109.2	8.2	1009	10	ADP38031	Adp38031 Synchroni
26	108	8.1	866	6	ABK65370	Abk65370 Arabidops
27	108	8.1	866	10	ADD30950	Add30950 Plant yie
28	108	8.1	866	12	ADB37148	Adb37148 Plant yie
29	108	8.1	866	12	AD141988	Adi41988 Plant tra
30	108	8.1	866	12	AD003516	Ado03516 Thalecres
31	108	8.1	1169	3	AAC47819	Aac47819 Arabidops
32	107.6	8.1	804	10	ADD29992	Add29992 Plant yie
33	107.6	8.1	804	12	AD144406	Adi44406 Plant tra
34	107	8.0	1277	13	ADRE0712	Adre0712 Cotton CD
35	106.4	8.0	349	3	AAC56698	Aac56698 Eucalyptu
36	106.4	8.0	828	12	AD000781	Ado00781 Rice Myb
37	106.4	8.0	1012	12	AD000778	Ado00778 Rice Myb
38	105.6	7.9	335	3	AAC56762	Aac56762 Eucalyptu
39	105	7.9	422	3	AAC56944	Aac56944 Pinus rad
40	105	7.9	436	3	AAC57318	Aac57318 Pinus rad
41	104.8	7.9	632	3	AAC56528	Aac56528 Eucalyptu
42	104.4	7.8	274	12	AD006568	Ado06568 Soybean t
43	103.8	7.8	589	13	ACN46091	Acn46091 Cotton pr
44	103.8	7.8	598	13	ACN54935	Acn54935 Cotton an
45	102.2	7.7	284	12	ADQ06567	Adq06567 Soybean t

ALIGNMENTS

RESULT 1
AD000777
ID AD000777 standard; cDNA; 1330 BP.
XX
AC AD000777;
XX
DT 26-AUG-2004 (first entry)
XX
DE Rice Myb protein OsmYBS1 cDNA.
XX
OS
XX
KW ss; gene; OsmYBS1; rice; plant; transgenic; Myb; transcription factor.
XX
OS Oryza sativa.
XX
FH Key Location/Qualifiers
FT CDS 20..940
FT /*tag= a
FT /product= "OsmYBS1"
XX
XX
XX US2004107456-A1.
XX
XX 03-JUN-2004.
XX
XX 30-JUL-2003; 2003US-00630636.
XX
XX 31-JUL-2002; 2002US-0399999P.
XX
XX (YUUS/) YU S.
XX
XX YU S;
XX
XX WPI: 2004-419531/39.
XX P-PSDB: AD000783.
XX
XX Novel Myb proteins derived from Oryza sativa and designated OsmYBS1,
XX OsmYBS2 or OsmYBS3; useful for regulating expression of a gene in plant
XX cell, preferably rice or barley cell.
XX
XX Claim 18; SEQ ID NO 1; 22pp; English.
XX
XX The invention relates to a pure polypeptide comprising a fully defined
XX Myb protein transcription factor OsmYBS1, OsmYBS2 or OsmYBS3 protein
XX sequence, where the polypeptide regulates expression of a gene in a cell.

CC The polypeptide is useful for regulating expression of a gene in a plant
 CC cell, preferably rice or barley cell. The method is useful for
 CC identifying a compound that modulates the activity of the polypeptide. An
 CC antibody to the polypeptide is useful for detecting the presence and
 CC distribution of OsMYBS1 proteins in tissues and in cellular compartments
 CC and for verifying the expression of OsMYBS proteins in a transgenic
 CC plant. Nucleic acid encoding the polypeptide is useful for generating a
 CC transgenic plant or producing to determine whether an OsMYBS mRNA is
 CC expressed in a tissue or cell. Nucleic acid encoding the polypeptide is
 CC also useful as primers in PCR-based detection methods, or as labelled
 CC probes in nucleic acid blots (e.g. Northern blots). The compound
 CC identified by the method is useful for enhancing or repressing the
 CC expression of a gene regulated by the polypeptide, e.g. alpha-amylase
 CC gene. The present sequence represents cDNA encoding the rice Myb protein
 CC OsMYBS1.

XX Sequence 1330 BP; 254 A; 438 C; 435 G; 203 T; 0 U; 0 Other;

Query Match 100.0%; Score 1330; DB 12; Length 1330;
 Best Local Similarity 100.0%; Pred. No. 3e-204;
 Matches 1330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGGATATCCACCAACCGATGACCTCCAGCGGCGACGACGACGACGCGCGGCG 60
 Db 1 GTGGGATATCCACCAACCGATGACCTCCAGCGGCGACGACGACGACGCGCGGCG 60
 QY 61 GCGCGCGCGGTGACACGAGGAGACGACAAAGCGCTTGAGAACGCGCTCGCGCTTGGC 120
 Db 61 GCGCGCGCGGTGACACGAGGAGACGACAAAGCGCTTGAGAACGCGCTCGCGCTTGGC 120
 QY 121 GCGCGCGCGCGCGCGCGGACGAGCGCGCGCGACGACGACTGTGCTCGCGCGCG 180
 Db 121 GCGCGCGCGCGCGCGCGGACGAGCGCGCGCGACGACGACTGTGCTCGCGCGCG 180
 QY 181 GAGCGTCCCGGCGGCGAGTCCGCGGAGAGGTGCGGAGGACATCAGAGGCGTGGTGA 240
 Db 181 GAGCGTCCCGGCGGCGAGTCCGCGGAGAGGTGCGGAGGACATCAGAGGCGTGGTGA 240
 QY 241 GAGCGTCCCGGCGGCGAGTCCGCGGAGAGGTGCGGAGGACATCAGAGGCGTGGTGA 300
 Db 241 GAGCGTCCCGGCGGCGAGTCCGCGGAGAGGTGCGGAGGACATCAGAGGCGTGGTGA 300
 QY 301 GCGCGCGCGCGCGCGGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
 Db 301 GCGCGCGCGCGCGCGGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
 QY 361 GCGCGAGAGCGCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
 Db 361 GCGCGAGAGCGCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
 QY 421 GAGCAGAGAGAGCGCAGAGGCGATCCCATGAGCGAGAGAGAGAGAGAGAGGCTTCTT 480
 Db 421 GAGCAGAGAGAGCGCAGAGGCGATCCCATGAGCGAGAGAGAGAGAGAGAGGCTTCTT 480
 QY 481 GGGGCTTGACAAAGTTCCGCAAGGGGACTGGCGGAGCATCTCGCGCAACTTCGTCATCTC 540
 Db 481 GGGGCTTGACAAAGTTCCGCAAGGGGACTGGCGGAGCATCTCGCGCAACTTCGTCATCTC 540
 QY 541 GCGGACCGCAACGAGGTGGGAGCGACGCGCAGAAAGTCTTCACTCGGCTCAACTCAT 600
 Db 541 GCGGACCGCAACGAGGTGGGAGCGACGCGCAGAAAGTCTTCACTCGGCTCAACTCAT 600
 QY 601 GAACCGGACCGCGCGCGCTTCAGCATCAAGCATCAACGAGCTCAACGCGCGGATCA 660
 Db 601 GAACCGGACCGCGCGCGCTTCAGCATCAAGCATCAACGAGCTCAACGCGCGGATCA 660
 QY 661 GGTGCGCGCGGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
 Db 661 GGTGCGCGCGGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
 QY 721 GCGCGCGCTGGCG 780
 Db 721 GCGCGCGCTGGCG 780

QY 781 GCCGCCATGCCCATGTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
 Db 781 GCCGCCATGCCCATGTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
 QY 841 GCCCGCGCGCGTGGACGACCGCGTGGTTCGCCCGCGCGCGCGCGCGCGCGCGCGCG 900
 Db 841 GCCCGCGCGCGTGGACGACCGCGTGGTTCGCCCGCGCGCGCGCGCGCGCGCGCGCG 900
 QY 901 CGTGGCTACCGCGCGCGCTCCGCGCAAGATGACCAATGACGCGCGCATGAGCGAGATGA 960
 Db 901 CGTGGCTACCGCGCGCGCTCCGCGCAAGATGACCAATGACGCGCGCATGAGCGAGATGA 960
 QY 961 GAGCAATTTCTTCT 1020
 Db 961 GAGCAATTTCTTCT 1020
 QY 1021 CGGCTCATGTCCT 1080
 Db 1021 CGGCTCATGTCCT 1080
 QY 1081 GATCAGATGCGCAAGAGTGCAGGAGTTGGGATTAAGCGAATAAGTAATTTTGTCT 1140
 Db 1081 GATCAGATGCGCAAGAGTGCAGGAGTTGGGATTAAGCGAATAAGTAATTTTGTCT 1140
 QY 1141 GACTGTTTGCAGTGTATCATCACTACACCGGTTAAAGCTTAGCTCCAAATGTGATGT 1200
 Db 1141 GACTGTTTGCAGTGTATCATCACTACACCGGTTAAAGCTTAGCTCCAAATGTGATGT 1200
 QY 1201 AATTACAGCGGCGCTCCGTAAGTGTGGCGCGCGATCGATGATCTTGAGGGTTCAT 1260
 Db 1201 AATTACAGCGGCGCTCCGTAAGTGTGGCGCGCGATCGATGATCTTGAGGGTTCAT 1260
 QY 1261 TAGGAGTTGATTTCCATTTTGTGATGTAATTTGGCAACTGTCTGATTTGACCAAAAA 1320
 Db 1261 TAGGAGTTGATTTCCATTTTGTGATGTAATTTGGCAACTGTCTGATTTGACCAAAAA 1320
 QY 1321 AAAAAAAAAA 1330
 Db 1321 AAAAAAAAAA 1330

RESULT 2
 ADO00780
 ID ADO00780 standard; cDNA; 918 BP.
 AC ADO00780;
 XX 26-AUG-2004 (first entry)
 XX Rice Myb protein OsMYBS1 CDS cDNA.
 XX s; gene; OsMYBS1; rice; plant; transgenic; Myb; transcription factor.
 XX Oryza sativa.
 XX Key Location/Qualifiers
 FT CDS 1..918
 FT /tag= a
 FT /partial
 FT /product= "OsMYBS1"
 FT /note= "No stop codon given"
 US2004107456-A1.
 PD 03-JUN-2004.
 XX 30-JUL-2003; 2003US-00630636.
 PF 31-JUL-2002; 2002US-0399999P.
 PR 31-JUL-2002; 2002US-0399999P.
 XX (YUSSI) YU S.
 XX

PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0136021P.
PR 25-MAY-1999; 99US-0136782P.
PR 27-MAY-1999; 99US-0136782P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138034P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139453P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 14-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 23-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.

PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157665P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159311P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 16-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.

PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 10.2%; Score 135.4; DB 3; Length 1161;
Best Local Similarity 70.4%; Pred. No. 8.1e-13;
Matches 181; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 393 ACGACGGCGGCAAGAGCTGCTCCAGGCGGAGAGAGAGGCGCAAGGCGCATCCATGGA 452
DB 419 ATGGAGGAAGAGTGGTGGCTCGAGAGCTGAGCAAGAGGAAGAAAGGAGTTCCATGGA 478
QY 453 CGGAGGAGAGAGAGGCTGCTTCTTGGGCGTGGAGCAAGTTGGCAAGGGGGACTGGC 512
DB 479 CTGAAGAGAGCATGCTGTTGTTCTTTGGGTTTGGACAGTTTGGAAAGGAGATTGGA 538
QY 513 GAGAGATCTCGCGCACTTCGTCATCTCGCGGACCGCAACGCGAGTGGCGAGCCAGCGGC 572
DB 539 GAGAGATTTCAAGAACTTTTGATCTCAAGAACTCCAACCAATTGCGAAGTCAATGCTC 598
QY 573 AGAAGTACTTTCATCCGCTCAATCTCATGAAACCGGACCGCGCTCCAGCATCCACG 632
DB 599 AAAAATACTTTCATCGGCTTAATCTGATGAACCGAGATGAAGGGGCTTGAATTCACG 658
QY 633 ACATCACCAGCGTCACG 649
DB 659 ACATCACCAGCTGTGAC 675

RESULT 4

ADP93595
ID ADP93595 standard; cDNA; 412 BP.

XX AC ADP93595;

XX 09-SEP-2004 (first entry)

XX Cotton expressed sequence tag, EST, #2606.

XX Cotton; ss; EST; expressed sequence tag; plant; plant protection;

KW plant improvement; marker-assisted breeding.

XX Gossypium hirsutum; variety Nucletron33B.

XX US2004123338-A1.

XX 24-JUN-2004.

XX 08-DEC-2000; 2000US-00732627.

XX 10-DEC-1999; 99US-0170255P.

XX (FINC/) FINCHER K L.

XX Fincher KL;

XX WPI; 2004-479807/45.

XX New substantially purified nucleic acid molecule that encodes a cotton
PT protein or its fragment, useful as molecular tool for the targeting and
isolation of novel genes for plant protection and improvement.

XX Claim 1; SEQ ID NO 2606; 30pp; English.

XX The invention relates to a substantially purified nucleic acid molecule

CC that encodes a cotton protein or its fragment comprising an EST
CC (expressed sequence tag) appearing as ADP90990-ADP95919. Also included
CC are a substantially purified cotton protein or its fragment encoded by a
CC nucleic acid molecule above and a transformed plant (having a nucleic
CC acid molecule which comprises: an exogenous promoter region which
CC functions in a plant cell to cause the production of a mRNA molecule; a
CC structural nucleic acid molecule comprising one of the ESTs or their
CC complements; a 3' non-translated sequence that functions in the plant
CC cell to cause termination of transcription and addition of polyadenylated
CC ribonucleotides to a 3' end of the mRNA molecule). The ESTs are useful as
CC molecular tool for the targeting and isolation of novel genes for plant
CC protection and improvement. The ESTs are useful for developing new
CC strategies for understanding critical plant developmental and metabolic
CC pathways, for isolating genes and promoters, for identifying and mapping
CC the genes involved in developmental and metabolic pathways, and for
CC determining gene function. The cotton nucleic acid molecules are useful
CC as molecular tags to isolate genetic regions, isolate genes, map genes,
CC and determine gene function. The nucleic acid molecules are useful for
CC determining if genes are members of a particular gene family and for use
CC in marker-assisted breeding programs. The present sequence is one of the
CC 4930 cotton ESTs of the invention. NOTE: The sequences are not displayed
CC in the specification but are available in electronic format from the
CC USPTO at seqdata.uspto.gov/sequence.html?docid=20040123338.

XX SQ Sequence 412 BP; 120 A; 83 C; 106 G; 102 T; 0 U; 1 Other;

Query Match 9.8%; Score 130.6; DB 12; Length 412;
Best Local Similarity 69.3%; Pred. No. 4.9e-12;
Matches 192; Conservative 0; Mismatches 84; Indels 1; Gaps 1;

QY 396 ACGGCGGCAAGAGTGTCTCAAGCGGAGAGAGGCGCAAGGCGCATCCATGAGCG 455
DB 125 ATGGAGGAAGAGAGTTGAGTCCAGCAAGAAAGAAAGAAAGAAATCCCATGAGCTG 184
QY 456 AGAAGAGACACAGCGCTGTTCTTGGGCTGAGCAAGTTGGCAAGGGGGAATGGCCGA 515
DB 185 AAGAAGAGCATAGGTTATTTCTACTGCTTTAGACAAGTTTGGAAAGAGATTGGAGAA 244
QY 516 GCATCTCGCGCACTTCGTCATCTCGCGGAGCCAGCGAGTGGCGAGCCAGCGCGCA 575
DB 245 GCATTTCAAGGAACTTTGTGATGATGAGAACTCCGACCGAGTGGCTAGCCATGCAGAA 304
QY 576 AGTACTTCATCCGCTCACTCATGAAACCGGACCGCGCGCTCCAGCATCCAGAC- 634
DB 305 AGTATTTTATACGCTTAATTCGATGAATAGAACCGCGCGGCTCCAGCATCCAGACN 364
QY 635 ATCACCAGCGTCACCGCGCGGATCAAGTCCGCGCGC 671
DB 365 ATCAGAGGTTAACATGCGGAGTCTCATCAAGC 401

RESULT 5

ADC4636
ID ADC4636 standard; DNA; 1432 BP.

XX AC ADC4636;

XX 18-DEC-2003 (first entry)

XX Thalecress transcription factor-like DNA G1362.

XX Thalecress; transcription factor-like protein; ds; seed trait;

KW transgenic; gene; plant size; stress tolerance; yield;

KW disease resistance; plant.

XX Arabidopsis thaliana.

XX US2003093837-A1.

XX 15-MAY-2003.

XX 01-NOV-2002; 2002US-00286264.

CC consecutive amino acids of any of the 55 250-500 residue amino acid
CC sequences (S1), given in the specification, altering the environmental
CC stress response or tolerance of a plant, or altering a plant's trait and
CC altering the expression levels of at least one gene in a plant. The
CC recombinant polynucleotide and methods are useful for altering the
CC regulation of gene expression of plants to modify the plant's traits, in
CC particular with respect to environmental stress responses (e.g. to viral
CC infection, fungal infection, microbial infection, herbicide resistance,
CC heat, cold, heavy metals, low light, drought, osmotic stress and salt
CC concentration). The present sequence is an environmental stress related
CC polynucleotide of the invention.

CC Sequence 1432 BP; 426 A; 269 C; 351 G; 386 T; 0 U; 0 Other;

CC Query Match 9.7%; Score 129.2; DB 10; Length 1432;

CC Best Local Similarity 70.3%; Pred. No. 8e-12; Indels 0; Gaps 0;

CC Matches 173; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Qy 413 TCCAAGCGGAGCAGAGAGCGCGCATCCCATGACGAGAGAGACAGCGCTG 472
Db 580 TCMAAGTCGATCAAGAGAGCGAAGGGTATCGGTGACAGAGATGAGACAGGTTA 639
Qy 473 TTCTTCTGGGCTGACAAAGTTCCGCAAGGGGCACTGGCGGACATCTCCGCAACTTC 532
Db 640 TTTCTTCTGGTGTGATTAAGTACGCGAAGGATGATGGCGTATTTCTCGCAACTTT 699
Qy 533 GTCAATCTGGCGAGCGCCCAACCGAGTGGCGAGCGGACGAGAGATTAATCTTATCCGCTC 592
Db 700 GTAGTAACAGAAACACGACCCCAAGTTGCGAGCCATGCTCAAAAGTATTTCAATTGCTTA 759
Qy 593 AACTCATGAACCGGACCGCGCGCTCCAGCATTCACGACATCACGACGCTCACCGCC 652
Db 760 AATTCAATGAACAAAGACAGAAAGGATCAAGCATTCACGACATCACTAGTGTGGCAAC 819
Qy 653 GGCGAT 658
Db 820 GCAGAT 825

RESULT 7
AD161458
ID AD161458 standard; cDNA; 1432 BP.

XX AD161458;

XX 22-APR-2004 (first entry)

DE cDNA encoding A. thaliana novel transcription factor (TF) #79.

XX Plant; transcription factor; TF; plant trait; gene; ss.

OS Arabidopsis thaliana.

XX US2003229915-A1.

XX 11-DEC-2003.

XX 22-NOV-2002; 2002US-00302267.

XX 18-FEB-1999; 99US-0120880P.

XX 22-FEB-1999; 99US-0121037P.

XX 11-MAR-1999; 99US-0124276P.

XX 15-APR-1999; 99US-0128450P.

XX 20-MAY-1999; 99US-0135134P.

XX 15-JUL-1999; 99US-0144153P.

XX 22-OCT-1999; 99US-0161143P.

XX 01-NOV-1999; 99US-0162656P.

XX 17-FEB-2000; 2000US-00506720.

PA (ADAM/) ADAM L.
PA (BROU/) BROU P.
PA (PINE/) PINEA O.
PA (REUB/) REUBER L.
PA (ZHAN/) ZHANG J.
PA (YONG/) YU G.
PA (JIANG/) JIANG C.
PA (SAMAH/) SAMAH R.
PA (PILG/) PILGRIM M.
PA (CREBE/) CREBELMAN R.

XX Keddle J., Fromm M., Heard J., Riechmann J.L., Adam L., Brown P.
PI Pineda O., Reuber L., Zhang J., Yu G., Jiang C., Samaha R., Pilgrim M.
PI Creelman R.

DR WPI: 2004-052052/05.
DR P-PSDB; AD161459.

DR New polynucleotide, useful in screening for a transcription factor that
modifies a plant trait.

PS Claim 1; SEQ ID NO 157; 16pp; English.

XX The present invention relates to the isolation of novel plant
CC (arabidopsis thaliana) polynucleotide sequences that encode transcription
CC factors (TFs), and the polypeptide sequences for the TFs. The
CC polynucleotide sequences are useful in screening for a transcription
CC factor that modifies a plant trait. Also disclosed is an expression
CC vector comprising a TF polynucleotide sequence, a host cell comprising
CC the expression vector, a transgenic plant comprising or ectopically
CC expressing an isolated TF polynucleotide sequence, a method for screening
CC for a molecule that modifies a plant trait, a method for producing a TF
CC transgenic plant, a method for identifying a sequence homologous to a TF
CC polynucleotide or polypeptide sequence, and a method for screening for a
CC transcription factor that modifies a plant trait. The present sequence
CC encodes a novel A. thaliana TF of the invention. Note: The sequence data
CC for this patent did not form part of the printed specification. The
CC complete sequence data for this patent was obtained in electronic format
CC directly from the USPTO web site at seqdata.uspto.gov.

XX Sequence 1432 BP; 426 A; 269 C; 351 G; 386 T; 0 U; 0 Other;

XX Query Match 9.7%; Score 129.2; DB 12; Length 1432;

XX Best Local Similarity 70.3%; Pred. No. 8e-12; Indels 0; Gaps 0;

XX Matches 173; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Qy 413 TCCAAGCGGAGCAGAGAGCGCGCATCCCATGACGAGAGAGACAGCGCTG 472
Db 580 TCMAAGTCGATCAAGAGAGCGAAGGGTATCGGTGACAGAGATGAGACAGGTTA 639
Qy 473 TTCTTCTGGGCTGACAAAGTTCCGCAAGGGGCACTGGCGGACATCTCCGCAACTTC 532
Db 640 TTTCTTCTGGTGTGATTAAGTACGCGAAGGATGATGGCGTATTTCTCGCAACTTT 699
Qy 533 GTCAATCTGGCGAGCGCCCAACCGAGTGGCGAGCGGACGAGAGATTAATCTTATCCGCTC 592
Db 700 GTAGTAACAGAAACACGACCCCAAGTTGCGAGCCATGCTCAAAAGTATTTCAATTGCTTA 759
Qy 593 AACTCATGAACCGGACCGCGCGCTCCAGCATTCACGACATCACGACGCTCACCGCC 652
Db 760 AATTCAATGAACAAAGACAGAAAGGATCAAGCATTCACGACATCACTAGTGTGGCAAC 819
Qy 653 GGCGAT 658
Db 820 GCAGAT 825

RESULT 8

AD002450
ID AD002450 standard; cDNA; 1432 BP.

XX AD002450;

XX

DT 01-JUL-2004 (first entry)
XX
XX Thalecress transcription factor CDNA #432.
XX Thalecress; transcription factor; ss; gene; plant; transgenic;
KW abiotic stress; cold tolerance; heat tolerance; drought; osmotic stress;
KW phosphate limitation; potassium limitation; nitrogen limitation;
KW hormone sensitivity; disease resistance; sugar sensing; seed germination;
KW flowering; inflorescence architectural change;
KW meristem cell differentiation; phyllotaxy; apical dominance;
KW trichome development; seed development; premature senescence;
KW delayed senescence; lethality; necrosis; plant size; leaf morphology;
KW seed morphology; secondary metabolism; light response; shade avoidance.
XX
XX Arabidopsis thaliana.
XX
XX US2004045049-A1.
XX
XX 04-MAR-2004.
XX
XX 10-APR-2003; 2003US-00412699.
XX
XX 13-SEP-1999; 99US-00394519.
XX 21-JAN-2000; 2000US-00489376.
XX 17-FEB-2000; 2000US-00506720.
XX 22-MAR-2000; 2000US-00532591.
XX 22-MAR-2000; 2000US-00533029.
XX 22-MAR-2000; 2000US-00533030.
XX 22-MAR-2000; 2000US-00533392.
XX 22-MAR-2000; 2000US-00533648.
XX 06-APR-2000; 2000WO-00509448.
XX 16-NOV-2000; 2000US-00713994.
XX 27-MAR-2001; 2001US-00819142.
XX 17-APR-2001; 2001US-00837444.
XX 30-JAN-2002; 2002US-00958131.
XX 14-JUN-2002; 2002US-00171468.
XX 09-AUG-2002; 2002US-00225066.
XX 09-AUG-2002; 2002US-00225067.
XX 09-AUG-2002; 2002US-00225068.
XX 17-DEC-2002; 2002US-0434166P.
XX 25-FEB-2003; 2003US-00374780.
XX
XX (ZHAN/) ZHANG J.
XX (FROM/) FROM M E.
XX (HEAR/) HEARD J E.
XX (RIEC/) RIECHMANN J L.
XX (ADAM/) ADAM L J.
XX (BROU/) BROUN P E.
XX (PINE/) PINEDA O.
XX (REUB/) REUBER T L.
XX (KEDD/) KEDDIE J S.
XX (YUGG/) YU G.
XX (JIAN/) JIANG C.
XX (SAMA/) SAMAH R S.
XX (PILG/) PILGRIM M L.
XX (CREE/) CREELMAN R A.
XX (DUBE/) DUBBEL A N.
XX (RATC/) RATCLIFFE O.
XX (KUMI/) KUMIMOTO R.
XX (SHER/) SHERMAN B K.
XX
XX Zhang J, Fromm ME, Heard JE, Riechmann JL, Adam LJ, Broun PE;
PI Pineda O, Reuber TL, Keddie JS, Yu G, Jiang C, Samaha RS;
PI Pilgrim ML, Creelman RA, Dubell AN, Ratcliffe O, Kumimoto R;
PI Sherman BK;
XX
XX WPI; 2004-225755/21.
XX P-PSDB; ADO002451.
XX
XX New transgenic plant, useful in developing phenotypes with altered or
PT improved characteristics or traits.
XX
XX Claim 1; SEQ ID NO 863; 213pp; English.

XX
CC The invention relates to a transgenic plant comprises a recombinant
CC polynucleotide having a polynucleotide sequence or its complementary
CC sequence comprising a sequence encoding a polypeptide, that initiates
CC transcription (i.e. a transcription factor) from Arabidopsis, Soybean,
CC Rice, Rape or Corn, comprising any of the sequences appearing as ADO01588
CC *ADO05527 or ADO03530-ADO03559. Also included are using a transgenic
CC plant to grow a progeny plant, an expression cassette (comprising a
CC constitutive, inducible or tissue-specific promoter and a recombinant
CC polynucleotide described above), a host cell comprising the expression
CC cassette, producing a modified plant having a modified trait, identifying
CC a factor that is modulated by or interacts with a polypeptide encoded by
CC the polynucleotide sequence and identifying at least one downstream
CC polynucleotide sequence that is subject to a regulatory effect of any of
CC the polypeptides encoded by the polynucleotide described above. The
CC transgenic plant is useful for producing a plant that has an altered
CC trait e.g. an enhanced tolerance to abiotic stress (increased tolerance
CC to chilling, germination in cold conditions freezing tolerance, tolerance
CC to heat, tolerance to drought, tolerance to osmotic stress, tolerance to
CC salt, tolerance to phosphate limitation, tolerance to potassium
CC limitation, decreased sensitivity to nitrogen limitation), altered
CC hormone sensitivity, reduced sensitivity to abscisic acid, an altered
CC response to ethylene, disease resistance, altered susceptibility to
CC Botrytis, altered susceptibility to Fusarium, altered susceptibility to
CC Erysiphe, altered susceptibility to Pseudomonas syringae, altered
CC susceptibility to Sclerotinia, altered sugar sensing, improved seed
CC germination and seedling vigor, early flowering, late flowering, extended
CC period of flowering, an inflorescence architectural change, a change in
CC stem bifurcations, a lack of a shoot meristem, reduced meristem cell
CC differentiation, altered phyllotaxy, altered branching pattern, reduced
CC apical dominance, reduced trichome density, ectopic trichome development,
CC altered trichome development, altered stem morphology, increased root
CC growth, increased root hairs, altered seed development, altered cell
CC proliferation/cell differentiation, premature senescence, delayed
CC senescence, lethality, increased necrosis, an increase in seedling or
CC plant size, decreased plant size, a change in leaf morphology, increased
CC altered leaf development, increased leaf size and mass, glossy leaves,
CC leaf cell expansion, change in seed morphology, altered seed coloration,
CC increased seed size, decreased seed size, altered seed shape, change in
CC leaf biochemistry, increased leaf wax, an alteration in leaf prenyl lipid
CC content, increased leaf insoluble sugars, decreased leaf insoluble
CC sugars, increased leaf anthocyanins, an alteration of leaf fatty acid
CC content, an alteration of leaf glucosinolate content, change in seed
CC biochemistry, an increase in seed oil content, decrease in seed oil
CC content, increase in seed fatty acid content, decrease in seed fatty acid
CC content, increase in seed protein content, decrease in seed protein
CC content, alteration in seed prenyl lipid content, increase in seed
CC sterols, upregulation of genes involved in secondary metabolism, increase
CC in root anthocyanins, increase in plant anthocyanins, and alteration in
CC light response or shade avoidance. The present sequence encodes a
CC thalecress transcription factor of the invention.
XX
SQ Sequence 1432 BP; 426 A; 269 C; 351 G; 386 T; 0 U; 0 Other;
Query Match 9.7%; Score 129.2; DB 12; Length 1432;
Best Local Similarity 70.3%; Pred. No. 8e-12;
Matches 173; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
Qy 413 TCCAGGCGGAGAGGAGGAGGCGCATCCATGACGAGAGGAGGACAGGCTG 472
Db 580 TCAAGTCCGATCAAGACACGAAAGGATGCGGTGACAGAAATGAGCAGAGTTA 639
Qy 473 TTCTTGCGGGGTGACAAATTGCGAGAGGGGACTGGGAGCATCTCGGCAACTTC 532
Db 640 TTCTTCTTGTTGGATGATGTAAGGGAAGTGATGGCGTAGCATTTCTCCCACTTT 639
Qy 533 GTCAATCTCGGAGACGCCAAGCGAGTGGCGAGCGCGGAGAAAGTAATCTCCGCTC 592
Db 700 GTAGTAACAAGAACACCGACCAAGTTGCGAGCGCATGCTCAAAAGTAATTCATTCGCTA 759
Qy 593 AACTCCATGAACCGGCGCGCGCGCTCCAGATCCAGACATCAACAGCATCAGCGCC 652
Db 760 AATTCAATGAACAAAGCAGAGGCGCTCAAGCATTCACGATCACTAATGTTGGCAAC 819

[illegible]

RESULT 11
ADQ03867
ID ADQ03867 standard; DNA; 257 BP.
XX AC
XX ADQ03867;
DT DT
DT 09-SEP-2004 (first entry)
XX DE
DE Maize transcription factor seqid 713.
XX
KW maize; soybean; Arabidopsis thaliana; transcription factor; homeobox;
KW Hhl1; leucine zipper; zinc finger; transformed plant;
KW metabolic pathway; mutation detection; polymorphism; plant trait;
KW genome mapping; gene identification; gene analysis; plant breeding;
KW transgenic; ds.
XX
OS Zea mays.
XX
PN US2004123339-A1.
XX PD
PD 24-JUN-2004.
XX PF
PF 06-AUG-2001; 2001US-00922293.
XX PR
PR 24-NOV-1997; 97US-0065700P.
PR 09-DEC-1997; 97US-0069472P.
PR 13-JAN-1998; 98US-0071479P.
PR 10-FEB-1998; 98US-0074201P.
PR 10-FEB-1998; 98US-0074280P.
PR 10-FEB-1998; 98US-0074281P.
PR 10-FEB-1998; 98US-0074282P.
PR 12-FEB-1998; 98US-0074565P.
PR 12-FEB-1998; 98US-0074566P.
PR 12-FEB-1998; 98US-0074567P.
PR 19-FEB-1998; 98US-0074789P.
PR 19-FEB-1998; 98US-0075459P.
PR 19-FEB-1998; 98US-0075460P.
PR 19-FEB-1998; 98US-0075461P.
PR 19-FEB-1998; 98US-0075462P.
PR 19-FEB-1998; 98US-0075463P.
PR 19-FEB-1998; 98US-0075464P.
PR 09-MAR-1998; 98US-0077229P.
PR 09-MAR-1998; 98US-0077230P.
PR 18-MAR-1998; 98US-0078368P.
PR 07-APR-1998; 98US-0080844P.
PR 27-APR-1998; 98US-0083067P.
PR 29-APR-1998; 98US-0083386P.

PR	29-APR-1998;	98US-0083387P.
PR	29-APR-1998;	98US-0083388P.
PR	29-APR-1998;	98US-0083389P.
PR	13-MAY-1998;	98US-0085222P.
PR	13-MAY-1998;	98US-0085223P.
PR	13-MAY-1998;	98US-0085224P.
PR	15-MAY-1998;	98US-0085533P.
PR	21-MAY-1998;	98US-0086183P.
PR	21-MAY-1998;	98US-0086184P.
PR	21-MAY-1998;	98US-0086185P.
PR	21-MAY-1998;	98US-0086186P.
PR	21-MAY-1998;	98US-0086187P.
PR	21-MAY-1998;	98US-0086188P.
PR	16-JUN-1998;	98US-0089524P.
PR	18-JUN-1998;	98US-0089793P.
PR	18-JUN-1998;	98US-0089806P.
PR	18-JUN-1998;	98US-0089807P.
PR	18-JUN-1998;	98US-0089808P.
PR	18-JUN-1998;	98US-0089810P.
PR	18-JUN-1998;	98US-0089811P.
PR	18-JUN-1998;	98US-0089812P.
PR	18-JUN-1998;	98US-0089813P.
PR	18-JUN-1998;	98US-0089814P.
PR	30-JUN-1998;	98US-0091247P.
PR	30-JUN-1998;	98US-0091405P.
PR	09-SEP-1998;	98US-0099667P.
PR	09-SEP-1998;	98US-0099668P.
PR	09-SEP-1998;	98US-0099670P.
PR	09-SEP-1998;	98US-0099697P.
PR	16-SEP-1998;	98US-0100672P.
PR	16-SEP-1998;	98US-0100673P.
PR	16-SEP-1998;	98US-0100674P.
PR	17-SEP-1998;	98US-0100963P.
PR	21-SEP-1998;	98US-0101130P.
PR	21-SEP-1998;	98US-0101131P.
PR	21-SEP-1998;	98US-0101132P.
PR	22-SEP-1998;	98US-0101343P.
PR	22-SEP-1998;	98US-0101344P.
PR	22-SEP-1998;	98US-0101347P.
PR	22-SEP-1998;	98US-0101508P.
PR	25-SEP-1998;	98US-0101707P.
PR	13-OCT-1998;	98US-0104124P.
PR	13-OCT-1998;	98US-0104126P.
PR	13-OCT-1998;	98US-0104127P.
PR	13-OCT-1998;	98US-0104128P.
PR	18-NOV-1998;	98US-0108996P.
PR	19-NOV-1998;	98US-0109018P.
PR	24-NOV-1998;	98US-0019912P.
PR	08-DEC-1998;	98US-0021029P.
PR	11-DEC-1998;	98US-0111981P.
PR	22-DEC-1998;	98US-0113244P.
PR	12-JAN-1999;	99US-00229413.
XX	(CONN/) CONNER T W.	
PA	(HECK/) HECK G R.	
PA	(LIUJ) LIU J.	
XX	Conner TW, Heck GR, Liu J;	
P1		
XX		
DR	WPI; 2004-468202/44.	
XX		
PT	New substantially purified nucleic acid molecule that encodes a maize,	
PT	soybean or Arabidopsis thaliana transcription factor or its fragment,	
PT	useful for genome mapping, gene identification and analysis or plant	
PT	breeding.	
XX		
PS	Claim 2; SEQ ID NO 713; 140bp; English.	
CC	The invention describes a substantially purified nucleic acid molecule	
CC	that encodes a maize, soybean or Arabidopsis thaliana transcription	
CC	factor or its fragment, where the maize or soybean transcription facto	
CC	r is homeobox, HNH, leucine zipper, zinc finger, or other transcription	
CC	factor. Also described are: a substantially purified maize or soybean	

transcription factor or its fragment defined above; a substantially purified antibody or its fragment which is capable of specifically binding to the transcription factor or its fragment above; a transformed plant; a method for determining a level or pattern in a plant cell of a transcription factor in a plant metabolic pathway; a method of determining a mutation in a plant; a method of producing a plant containing an overexpressed protein or reduced levels of plant transcription factor; a method of determining an association between a polymorphism and a plant trait; and a method of isolating a nucleic acid that encodes a plant transcription factor or its fragment. The nucleic acid molecules, proteins and their fragments are useful for genome mapping, gene identification and analysis, plant breeding, preparation of constructs for use in plant gene expression and transgenic plants. The nucleic acid molecules are useful as markers or probes. This sequence represents a maize transcription factor polynucleotide.

Sequence 257 BP; 70 A; 57 C; 87 G; 43 T; 0 U; 0 Other;

Query Match 9.3%; Score 123.2; DB 12; Length 257;
Best Local Similarity 74.6%; Pred. No. 7.5e-11;
Matches 182; Conservative 0; Mismatches 58; Indels 4; Gaps 2;

Qy 375 AGGCGCGGCGCGCGGTACGACGCGCGCAAGCTCTCCAG---GCGAGCAGAGGA 431
Db 15 AGAGTGGCGAGAGAGACCCATGGGCAAGGGGTCCGCAAGTCCGCGAGGAGC 74
Qy 432 GCGCGAAGGGGATCCCATGACGAGAGAGACAGAGCTGTCTTCTGGCGGCTGACA 491
Db 75 GTAGGAGAGGATCGCCCTGACAGAGAGACAGACAGATTTCTCTTGGACTTGA 134
Qy 492 AGTTGGGCAAGGGGACTGGCGAGACATTCGCGCACTTCATCTCGCGAGCCAA 551
Db 135 AGTATGGCAAGGTGACTGGCGAGCATCTCGCG--AACTTGTGATCTCAAGAGCCGA 193
Qy 552 CGCAGGTGGCGAGCGACCGGAGAGAGTACTTATCCGCTCACTCCATGAACCGCCAGC 611
Db 194 CTCAAGTAGCTAGTCTATGACAGAGATCTTATCCGCTGAACTGAACAGAGAAA 253
Qy 612 GCCG 615
Db 254 GGCG 257

RESULT 12
ADQ04396
ID ADQ04396 standard; DNA; 386 BP.

XX AC ADQ04396;

DT 09-SEP-2004 (first entry)

XX DE Maize transcription factor seqid 1242.

XX maize; soybean; Arabidopsis thaliana; transcription factor; homeobox;
XX HbH; leucine zipper; zinc finger; transformed plant; plant;
XX metabolic pathway; mutation detection; polymorphism; plant trait;
XX genome mapping; gene identification; gene analysis; plant breeding;
XX transgenic; de.

XX Zea mays.

XX US2004123339-A1.

XX PD 24-JUN-2004.

XX PF 06-AUG-2001; 2001US-00922293.

XX PR 24-NOV-1997; 97US-0067000P.

XX PR 09-DEC-1997; 97US-0069472P.

XX PR 13-JAN-1998; 98US-0071479P.

XX PR 10-FEB-1998; 98US-0074201P.

XX PR 10-FEB-1998; 98US-0074280P.

XX PR 10-FEB-1998; 98US-0074281P.

PR 10-FEB-1998; 98US-0074282P.
PR 12-FEB-1998; 98US-0074565P.
PR 12-FEB-1998; 98US-0074566P.
PR 12-FEB-1998; 98US-0074567P.
PR 19-FEB-1998; 98US-0074789P.
PR 19-FEB-1998; 98US-0075459P.
PR 19-FEB-1998; 98US-0075460P.
PR 19-FEB-1998; 98US-0075461P.
PR 19-FEB-1998; 98US-0075462P.
PR 19-FEB-1998; 98US-0075463P.
PR 19-FEB-1998; 98US-0075464P.
PR 09-MAR-1998; 98US-0077229P.
PR 09-MAR-1998; 98US-0077230P.
PR 09-MAR-1998; 98US-0077231P.
PR 18-MAR-1998; 98US-0078368P.
PR 07-APR-1998; 98US-0080844P.
PR 27-APR-1998; 98US-0083067P.
PR 29-APR-1998; 98US-0083386P.
PR 29-APR-1998; 98US-0083387P.
PR 29-APR-1998; 98US-0083388P.
PR 29-APR-1998; 98US-0083389P.
PR 13-MAY-1998; 98US-0085222P.
PR 13-MAY-1998; 98US-0085223P.
PR 13-MAY-1998; 98US-0085224P.
PR 15-MAY-1998; 98US-0085333P.
PR 21-MAY-1998; 98US-0086183P.
PR 21-MAY-1998; 98US-0086184P.
PR 21-MAY-1998; 98US-0086185P.
PR 21-MAY-1998; 98US-0086186P.
PR 21-MAY-1998; 98US-0086187P.
PR 21-MAY-1998; 98US-0086188P.
PR 16-JUN-1998; 98US-0089524P.
PR 18-JUN-1998; 98US-0089793P.
PR 18-JUN-1998; 98US-0089806P.
PR 18-JUN-1998; 98US-0089807P.
PR 18-JUN-1998; 98US-0089808P.
PR 18-JUN-1998; 98US-0089810P.
PR 18-JUN-1998; 98US-0089811P.
PR 18-JUN-1998; 98US-0089812P.
PR 18-JUN-1998; 98US-0089813P.
PR 18-JUN-1998; 98US-0089814P.
PR 30-JUN-1998; 98US-0091247P.
PR 30-JUN-1998; 98US-0091248P.
PR 30-JUN-1998; 98US-0091405P.
PR 09-SEP-1998; 98US-0093668P.
PR 09-SEP-1998; 98US-0093669P.
PR 09-SEP-1998; 98US-0093670P.
PR 09-SEP-1998; 98US-0093671P.
PR 16-SEP-1998; 98US-0100672P.
PR 16-SEP-1998; 98US-0100673P.
PR 16-SEP-1998; 98US-0100674P.
PR 17-SEP-1998; 98US-0100963P.
PR 21-SEP-1998; 98US-0101130P.
PR 21-SEP-1998; 98US-0101131P.
PR 21-SEP-1998; 98US-0101132P.
PR 22-SEP-1998; 98US-0101343P.
PR 22-SEP-1998; 98US-0101344P.
PR 22-SEP-1998; 98US-0101347P.
PR 22-SEP-1998; 98US-0101508P.
PR 25-SEP-1998; 98US-0101707P.
PR 13-OCT-1998; 98US-0104124P.
PR 13-OCT-1998; 98US-0104126P.
PR 13-OCT-1998; 98US-0104127P.
PR 13-OCT-1998; 98US-0104128P.
PR 18-NOV-1998; 98US-0108966P.
PR 19-NOV-1998; 98US-0109018P.
PR 24-NOV-1998; 98US-0109129P.
PR 08-DEC-1998; 98US-00210297.
PR 11-DEC-1998; 98US-0111981P.
PR 22-DEC-1998; 98US-0113224P.
PR 12-JAN-1999; 99US-00229413.

XX PA (CONV/) CONNER T W.
XX PA (HECK/) HECK G R.

PR 13-OCT-1998; 98US-0104124P.
 PR 13-OCT-1998; 98US-0104126P.
 PR 13-OCT-1998; 98US-0104127P.
 PR 13-OCT-1998; 98US-0104128P.
 PR 18-NOV-1998; 98US-0108996P.
 PR 19-NOV-1998; 98US-0109018P.
 PR 24-NOV-1998; 98US-00199129.
 PR 08-DEC-1998; 98US-00210297.
 PR 11-DEC-1998; 98US-0111981P.
 PR 22-DEC-1998; 98US-0113224P.
 PR 12-JAN-1999; 98US-00229413.
 XX
 PA (CONN/) CONNER T W.
 PA (HECK/) HECK G R.
 PA (LIU/) LIU J.
 XX
 PI Conner TW, Heck GR, Liu J;
 XX
 DR WPI; 2004-468202/44.
 XX
 PT New substantially purified nucleic acid molecule that encodes a maize,
 PT soybean or Arabidopsis thaliana transcription factor or its fragment,
 PT useful for genome mapping, gene identification and analysis or plant
 PT breeding.
 XX
 PS Claim 2; SEQ ID NO 618; 1400p; English.
 XX
 CC The invention describes a substantially purified nucleic acid molecule
 CC that encodes a maize, soybean or Arabidopsis thaliana transcription
 CC factor or its fragment, where the maize or soybean transcription factor
 CC is homeobox, HLH, leucine zipper, zinc finger, or other transcription
 CC factor. Also described are: a substantially purified maize or soybean
 CC transcription factor or its fragment defined above; a substantially
 CC purified antibody or its fragment which is capable of specifically
 CC binding to the transcription factor or its fragment above; a transformed
 CC plant; a method for determining a level or pattern in a plant cell of a
 CC transcription factor in a plant metabolic pathway; a method of
 CC determining a mutation in a plant; a method of producing a plant
 CC containing an overexpressed protein ore reduced levels of plant
 CC transcription factor; a method of determining an association between a
 CC polymorphism and a plant trait; and a method of isolating a nucleic acid
 CC that encodes a plant transcription factor or its fragment. The nucleic
 CC acid molecule, proteins and their fragments are useful for genome
 CC mapping, gene identification and analysis, plant breeding, preparation of
 CC constructs for use in plant gene expression and transgenic plants. The
 CC nucleic acid molecules are useful as markers or probes. This sequence
 CC represents a maize transcription factor polynucleotide.
 CC
 XX
 SQ Sequence 315 BP; 85 A; 81 C; 86 G; 63 T; 0 U; 0 Other;
 Query Match 9.0%; Score 119.8; DB 12; Length 315;
 Best Local Similarity 77.5%; Pred. No. 2.6e-10;
 Matches 145; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

ID AD003861 standard; DNA; 144 BP.
 XX
 AC AD003861;
 XX
 DT 09-SEP-2004 (first entry)
 XX
 DE Maize transcription factor seqid 707.
 XX
 KW maize; soybean; Arabidopsis thaliana; transcription factor; homeobox;
 KW HLH; leucine zipper; zinc finger; transformed plant; plant;
 KW metabolic pathway; mutation detection; polymorphism; plant trait;
 KW genome mapping; gene identification; gene analysis; plant breeding;
 KW transgenic; ds.
 XX
 OS Zea mays.
 XX
 PN US2004123339-A1.
 XX
 PD 24-JUN-2004.
 XX
 PF 06-AUG-2001; 2001US-00922293.
 XX
 PR 24-NOV-1997; 97US-0067000P.
 PR 09-DEC-1997; 97US-0069472P.
 PR 13-JAN-1998; 98US-0071479P.
 PR 10-FEB-1998; 98US-0074201P.
 PR 10-FEB-1998; 98US-0074280P.
 PR 10-FEB-1998; 98US-0074281P.
 PR 10-FEB-1998; 98US-0074282P.
 PR 12-FEB-1998; 98US-0074565P.
 PR 12-FEB-1998; 98US-0074566P.
 PR 12-FEB-1998; 98US-0074567P.
 PR 19-FEB-1998; 98US-0074789P.
 PR 19-FEB-1998; 98US-0075459P.
 PR 19-FEB-1998; 98US-0075460P.
 PR 19-FEB-1998; 98US-0075461P.
 PR 19-FEB-1998; 98US-0075462P.
 PR 19-FEB-1998; 98US-0075463P.
 PR 19-FEB-1998; 98US-0075464P.
 PR 09-MAR-1998; 98US-0077228P.
 PR 09-MAR-1998; 98US-0077230P.
 PR 09-MAR-1998; 98US-0077231P.
 PR 18-MAR-1998; 98US-0078368P.
 PR 07-APR-1998; 98US-0080844P.
 PR 27-APR-1998; 98US-0083067P.
 PR 29-APR-1998; 98US-0083386P.
 PR 29-APR-1998; 98US-0083387P.
 PR 29-APR-1998; 98US-0083388P.
 PR 29-APR-1998; 98US-0083389P.
 PR 13-MAY-1998; 98US-0085222P.
 PR 13-MAY-1998; 98US-0085223P.
 PR 13-MAY-1998; 98US-0085224P.
 PR 15-MAY-1998; 98US-0085533P.
 PR 21-MAY-1998; 98US-0086183P.
 PR 21-MAY-1998; 98US-0086184P.
 PR 21-MAY-1998; 98US-0086185P.
 PR 21-MAY-1998; 98US-0086186P.
 PR 21-MAY-1998; 98US-0086187P.
 PR 21-MAY-1998; 98US-0086188P.
 PR 16-JUN-1998; 98US-0089524P.
 PR 18-JUN-1998; 98US-0089529P.
 PR 18-JUN-1998; 98US-0089806P.
 PR 18-JUN-1998; 98US-0089807P.
 PR 18-JUN-1998; 98US-0089808P.
 PR 18-JUN-1998; 98US-0089810P.
 PR 18-JUN-1998; 98US-0089811P.
 PR 18-JUN-1998; 98US-0089812P.
 PR 18-JUN-1998; 98US-0089813P.
 PR 18-JUN-1998; 98US-0089814P.
 PR 30-JUN-1998; 98US-0091247P.
 PR 30-JUN-1998; 98US-0091405P.
 PR 09-SEP-1998; 98US-0099667P.
 PR 09-SEP-1998; 98US-0099668P.

PR 09-SEP-1998; 98US-0099670P.
PR 09-SEP-1998; 98US-0099671P.
PR 16-SEP-1998; 98US-0100672P.
PR 16-SEP-1998; 98US-0100673P.
PR 16-SEP-1998; 98US-0100674P.
PR 17-SEP-1998; 98US-0100963P.
PR 21-SEP-1998; 98US-0101131P.
PR 21-SEP-1998; 98US-0101132P.
PR 21-SEP-1998; 98US-0101133P.
PR 22-SEP-1998; 98US-0101344P.
PR 22-SEP-1998; 98US-0101347P.
PR 22-SEP-1998; 98US-0101508P.
PR 23-SEP-1998; 98US-0101707P.
PR 13-OCT-1998; 98US-0104124P.
PR 13-OCT-1998; 98US-0104126P.
PR 13-OCT-1998; 98US-0104127P.
PR 13-OCT-1998; 98US-0104128P.
PR 18-NOV-1998; 98US-0108996P.
PR 19-NOV-1998; 98US-0109018P.
PR 24-NOV-1998; 98US-0019129.
PR 08-DEC-1998; 98US-00210297.
PR 11-DEC-1998; 98US-0111981P.
PR 22-DEC-1998; 98US-0113224P.
PR 12-JAN-1999; 99US-00229413.

XX (CONN/) CONNER T W.
PA (HECK/) HECK G R.
PA (LIU/) LIU J.

XX Conner TW, Heck GR, Liu J;

XX WPI; 2004-468202/44.

PT New substantially purified nucleic acid molecule that encodes a maize,
PT soybean or Arabidopsis thaliana transcription factor or its fragment,
PT useful for genome mapping, gene identification and analysis or plant
PT breeding.

PS Claim 2; SEQ ID NO 707; 140bp; English.

XX The invention describes a substantially purified nucleic acid molecule
CC that encodes a maize, soybean or Arabidopsis thaliana transcription
CC factor or its fragment, where the maize or soybean transcription factor
CC is homeobox, HLF, leucine zipper, zinc finger, or other transcription
CC factor. Also described are: a substantially purified maize or soybean
CC transcription factor or its fragment defined above; a substantially
CC purified antibody or its fragment which is capable of specifically
CC binding to the transcription factor or its fragment above; a transformed
CC plant; a method for determining a level or pattern in a plant cell of a
CC transcription factor in a plant metabolic pathway; a method of
CC determining a mutation in a plant; a method of producing a plant
CC containing an overexpressed protein or reduced levels of plant
CC transcription factor; a method of determining an association between a
CC polymorphism and a plant trait; and a method of isolating a nucleic acid
CC that encodes a plant transcription factor or its fragment. The nucleic acid
CC molecules, proteins and their fragments are useful for genome
CC mapping, gene identification and analysis, plant breeding, preparation of
CC constructs for use in plant gene expression and transgenic plants. The
CC nucleic acid molecules are useful as markers or probes. This sequence
CC represents a maize transcription factor polynucleotide.

XX Sequence 144 BP; 27 A; 39 C; 54 G; 24 T; 0 U; 0 Other;

XX Query Match 8.9%; Score 119; DB 12; Length 144;

XX Best Local Similarity 89.5%; Pred. No. 3.6e-10;

XX Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 428 GAGAGGCGCAAGGCGATCCATGACGAGGAGAGAGACAGAGCTGTTCTTGCGGGCTG 487
DB 2 GAGGCGGCGCAAGGCGATCCGTCGACGGAAGAGAGACAGAGCTGTTCTTGCGAGCTG 61
QY 488 GACAAGTTGGCAAGGGGAGCTGGCGAGCATCTTCGGCAACTTCGTACTCTCGGAGC 547

DB 62 GACAAGTTCTGCAAGGGCGAGCTGCGAGCATCTCGCAACTTCGTACTCTCGGAGC 121
QY 548 CCAAGCGAGGTGGCGAGCCACGC 570
DB 122 CCGACGAGGTGGCGAGCCACGC 144

RESULT 15
ABK65275
ID ABK65275 standard; cDNA; 867 BP.

AC ABK65275;

DT 02-JUL-2002 (first entry)

XX Arabidopsis cDNA encoding a transcription factor #127.

KM Plant, ss; gene; transcription factor; transgenic; agriculture;
KM metabolic chemical; environmental stress; drought;
KM microbial disease resistance; herbicide resistance; seed yield;
KM fruit yield; growth rate; leaf senescence; flower senescence.

OS Arabidopsis thaliana.

XX WO200215675-A1.

XX 28-FEB-2002.

PF 22-AUG-2001; 2001WO-US026189.

XX 22-AUG-2000; 2000US-0227439P.

PR 16-NOV-2000; 2000US-00713994.

PR 18-APR-2001; 2001US-00837944.

XX (MEND-) MENDEL BIOTECHNOLOGY INC.

PA (PIUG/) PILGRIM M.

PA (CREE/) CREELMAN R.

PA (DUBE/) DUBIELL A J.

PA (HEAR/) HEARD J.

PA (JIANG/) JIANG C.

PA (KEDD/) KEDDIE J.

PA (ADAM/) ADAM L.

PA (RATC/) RATCLIFF O.

PA (REUB/) REUBER J L.

PA (RIEC/) RIECHMANN J L.

PA (YUGG/) YU G.

PA (PINED/) PINEDA O.

XX Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddie J,

PI Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;

XX WPI; 2002-292022/33.

DR P-PSDB; AAU93089.

XX An isolated or recombinant polynucleotide used to produce a transgenic

PT plant.

XX Claim 4; Page 527-528; 941bp; English.

XX The invention relates to 1 of 232 isolated or recombinant polynucleotides
CC encoding an Arabidopsis thaliana transcription factor, their variants,
CC complements, fragments, or related polynucleotide with 31% to 95%
CC sequence identity, where the plant possesses an altered trait as compared
CC to a wild-type or reference plant, or the plant exhibits an altered
CC phenotype as compared to a wild-type or reference plant, or the plant
CC exhibits ectopic expression or altered expression of one or more genes
CC associated with a plant trait as compared to a wild plant. Also included
CC are a transgenic plant comprising the polynucleotides, a computer
CC readable medium having stored sequence information, and identifying a
CC homologue sequence from a database comprising a plurality of known plant
CC sequences comprising inputting sequence information selected from one of
CC 464 fully defined sequences given in the specification. The isolated or

This Page Blank (uspto)

QY 61 GGGCGCGCGCTGACCAAGGAGACGACGAGCGTTGAGAACCGCGCTCGCGGCTTGGCG 120
DB 61 GCGCGCGCGCTGAGACAGGAGAGAGACGACGAGCGTTGAGAACCGCGCTCGCGGCTTGGCG 120
QY 121 GGGCGCGCGCGCGCGAGCGAGCG 180
DB 121 GGGCGCGCGCGCGCGAGCGAGCG 180
QY 181 GAGCGTCCCGGCGCGAGGTCGCGAGAGAGTCGAGAGCGACCTAAGAGCGCTGAGGA 240
DB 181 GAGCGTCCCGGCGCGAGGTCGCGAGAGAGTCGAGAGCGACCTAAGAGCGCTGAGGA 240
QY 241 GAGCGTCCCGGCGCGACG 300
DB 241 GAGCGTCCCGGCGCGACG 300
QY 301 GCGCGCGCGCGCGCGAGCGAGCG 360
DB 301 GCGCGCGCGCGCGCGAGCGAGCG 360
QY 361 GCGCGAGAGAGCGCAAGGCG 420
DB 361 GCGCGAGAGAGCGCAAGGCG 420
QY 421 GAGAGAGAGAGAGCG 480
DB 421 GAGAGAGAGAGAGCG 480
QY 481 GGGCGTGGACAAAGTTCCGCGCAAGGGGGGACTGGCGGAGATCTCGCGCAACTTCGTCATCTC 540
DB 481 GGGCGTGGACAAAGTTCCGCGCAAGGGGGGACTGGCGGAGATCTCGCGCAACTTCGTCATCTC 540
QY 541 GCGGACCGCGCAACGAGGTGGCGAGCCACGCGCAGAAAGTACTTCTCGCGCTCAACTCCAT 600
DB 541 GCGGACCGCGCAACGAGGTGGCGAGCCACGCGCAGAAAGTACTTCTCGCGCTCAACTCCAT 600
QY 601 GAACCGGAGCGCGCGCGCTTCAGCATTCACGACATCAACGAGTCAACCGCGCGCGCATCA 660
DB 601 GAACCGGAGCGCGCGCGCTTCAGCATTCACGACATCAACGAGTCAACCGCGCGCGCATCA 660
QY 661 GGTGCGCGCGAGAGGGGGCGCGCGATCAACGCGCAACGAGCGCGCGCGCGCGCGCGCGCG 720
DB 661 GGTGCGCGCGAGAGGGGGCGCGCGATCAACGCGCAACGAGCGCGCGCGCGCGCGCGCGCG 720
QY 721 GCGCGCGCTGGCGCGCGCGCGCGCATGAAGACCAACGAGCGCGCGCGCGCGCGCGCGCGCG 780
DB 721 GCGCGCGCTGGCGCGCGCGCGCGCATGAAGACCAACGAGCGCGCGCGCGCGCGCGCGCGCG 780
QY 781 GCGCGCGCATGCGCATGTAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCAT 840
DB 781 GCGCGCGCATGCGCATGTAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCAT 840
QY 841 GCGCGCGCGCGTGGCGAGCGCGCGGTGTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
DB 841 GCGCGCGCGCGTGGCGAGCGCGCGGTGTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
QY 901 CGTGGGCTAACCGCGCGCGCGCGCGCAAGATGCAACCAATGACGCGCGCGCGCGCGCGCGCG 960
DB 901 CGTGGGCTAACCGCGCGCGCGCGCGCAAGATGCAACCAATGACGCGCGCGCGCGCGCGCGCG 960
QY 961 GCAGCAATTTCTTCT 1020
DB 961 GCAGCAATTTCTTCT 1020
QY 1021 CGGCTCATCGCTCCGCTGATCACTTGTCTCTCACTCACTCACTCACTCACTCACTCACTCACT 1080
DB 1021 CGGCTCATCGCTCCGCTGATCACTTGTCTCTCACTCACTCACTCACTCACTCACTCACTCACT 1080
QY 1081 GATCAGATGCGCAAGAGTGGAGGATTAAGCGCAATAAGTAAGTATTTTGGT 1140
DB 1081 GATCAGATGCGCAAGAGTGGAGGATTAAGCGCAATAAGTAAGTATTTTGGT 1140

QY 1141 GACTGTTGGCAAGTATCATACGTAACCCCGGTGAAGCTTAGTCCAAATGTGATGT 1200
DB 1141 GACTGTTGGCAAGTATCATACGTAACCCCGGTGAAGCTTAGTCCAAATGTGATGT 1200
QY 1201 AATTAGCAGCGGCTTCCGTAAGTGGTGGCGCGCATGATGATCTTSCAGAGGTTGCAAT 1260
DB 1201 AATTAGCAGCGGCTTCCGTAAGTGGTGGCGCGCATGATGATCTTSCAGAGGTTGCAAT 1260
QY 1261 TAGGATTTGATTTTCATTTTGTGATGTAATTTTCCCACTGTCTCATTTGACCAAAAAA 1320
DB 1261 TAGGATTTGATTTTCATTTTGTGATGTAATTTTCCCACTGTCTCATTTGACCAAAAAA 1320
QY 1321 AAAAAAAAAA 1330
DB 1321 AAAAAAAAAA 1330

RESULT 2

US-10-437-963-59484/c
; Sequence 59484, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barabazuk, Brad
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ. ID NOS: 204966
; SEQ. ID NO 59484
; LENGTH: 1433
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_61102C.1
US-10-437-963-59484

Query Match 97.3%; Score 1293.8; DB 19; Length 1433;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1298; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 20 ATGACCTCCAGAGCGCGAG 79
DB 1433 ATGACCTCCAGAGCGCGAG 1374
QY 80 GAGGACGACAAAGGCGTTTCGAGAAAGCGCTGCGCGCTTTCGCGCGCGCGCGCGCGCGCGCGAG 139
DB 1373 GAGGACGACAAAGGCGTTTCGAGAAAGCGCTGCGCGCTTTCGCGCGCGCGCGCGCGCGCGAG 1314
QY 140 GAGAGCGCGCGCGAG 199
DB 1313 GAGAGCGCGCGCGAG 1254
QY 200 TCGCGGAGAGAGAGTGGCGAGGCGCATAGAGAGCGCTGTGAGAGAGCTTCGCGCGCATTCGAC 259
DB 1253 TCGCGGAGAGAGAGTGGCGAGGCGCATAGAGAGCGCTGTGAGAGAGCTTCGCGCGCATTCGAC 1194
QY 260 GCGGCGCGCGTTCGCGCTCCGCGCGCGCTACGCGCGGAGAGAGTCCGCGCGCGCGCGCGAGCGA 319
DB 1193 GCGGCGCGCGTTCGCGCTCCGCGCGCGCTACGCGCGGAGAGAGTCCGCGCGCGCGCGCGAGCGA 1134
QY 320 GCGGAGCGCGCGCGCGCGCGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 379
DB 1133 GCGGAGCGCGCGCGCGCGCGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1074
QY 380 GCGGCGCGCGGTACGAGCGCGCGAGAGAGTCTCTCCAGGCGGAGAGAGAGAGAGAGAGAG 439

Db 841 CCGGTGCTGTTCCCGCGGCGCACGGCGGTAGTCTGTCGCCCTGACCTACCGGCGCT 900
QY 920 CCGGCCAAGATGACCA 937
Db 901 CCGGCCAAGATGACCA 918

RESULT 4

US-10-425-115-168067
; Sequence 168067, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 168067
; LENGTH: 1299
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_84858C.1
US-10-425-115-168067

Query Match 36.0%; Score 479.2; DB 20; Length 1299;
Best Local Similarity 75.0%; Pred. No. 2.9e-119;
Matches 769; Conservative 0; Mismatches 193; Indels 64; Gaps 11;

QY 3 GCGAGATCCACACCCGATGACTCCGAGGCGGAGACGACGACCAAGCGGCGCGG 62
Db 40 GAGAGACCGCCCAACCCGCGGACCAAGACATCTGGAGTACACTCTGCGACGAC 99
QY 63 CCGGCGGCTGACACGAGGAGACGACAGGCGTTGAGAAAGCGCTCGCGCTTGC 122
Db 100 CCGGCGGCTGACACGAGGAGACGAGGCGTTGAGAAAGCGCTCGCGCTTGC 159
QY 123 CCGCGCGCGCGGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 182
Db 160 CCGCGCGCGCGGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 213
QY 183 GCGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 242
Db 214 GCGTGC---GGCGGAAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 270
QY 243 ACCTGCGGCGCATGACGCGGCGCGCTCCGCTCCGCGCTACGCGGAGAGG--GT 299
Db 271 ACCTGCGGCGCATGACGCGGCGCGCTCCGCTCCGCGCTACGCGGAGAGG--GT 330
QY 300 CCGGCGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 350
Db 331 CCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 390
QY 351 -----GCGACACCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 404
Db 391 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 447
QY 405 AGAGCTCTCCAAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 464
Db 448 AGAGCTCTCCAAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 507
QY 465 ACAGGCTGTTTCTGCTGGGCTGAGCAAGTTCCGCAAGGGGAGATTGGCGAGCATCTCCG 524
Db 508 ACAGGCTGTTTCTGCTGGGCTGAGCAAGTTCCGCAAGGGGAGATTGGCGAGCATCTCCG 567
QY 525 GCAACTTCGATCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 584
Db 568 GCAACTTCGATCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 627

QY 585 TCCGCTCAATCTCATGAACCGGAGACCGCGCGCTTCAGATTCAGACATCAACAGCG 644
Db 628 TCCGCTCAATCTCATGAACCGGAGACCGGAGACCGCGCTTCAGATTCAGACATCAACAGCG 687
QY 645 TCAACCGCGGAGTCAAGTCCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 704
Db 688 TCAACCGCGGAGTCAAGTCCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 747
QY 705 CCGGCAACCCCGGCGGCGCGCTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 764
Db 748 CCGG-----GGGATGTCGATGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 792
QY 765 ACCCGGCG 824
Db 793 -----CGGTTCCCGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
QY 825 TCGCGCGCGCATGATG---GCGCGCGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 880
Db 841 TCGCGCGCGCATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
QY 881 --CACGCGCGTACGTCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 938
Db 901 ATCACTGCGCTATGTCCTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
QY 939 GACGCGCATGACGCG-ACATGACAGCATTTTCTCTCTCTCTCTCTCTCTCTCTCTCT 997
Db 961 GATGCTTACAGACCGGATTTGCGCATTCGCGAGCTGACATCCCTTGTCTCTCTCTCTCT 1020
QY 998 TGATTT 1003
Db 1021 TGATTT 1026

RESULT 5

US-10-739-930-2650
; Sequence 2650, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(5337)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 2650
; LENGTH: 1516
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: ZEAMA-23APR03-CLUSTER42_46
US-10-739-930-2650

Query Match 33.5%; Score 445.4; DB 20; Length 1516;
Best Local Similarity 76.8%; Pred. No. 4.3e-110;
Matches 732; Conservative 1; Mismatches 157; Indels 63; Gaps 13;

QY 28 CCAAGCGCGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 86
Db 282 CCAAGCATCTGCAATGACACTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 341
QY 87 ACAAGGCGTTG-AGAACGCGCTCGCGCTTGGCGGCGCGCGCGCGCGCGCGCGCGAGC 145
Db 342 ACAAGGCGTTGAGAAACGCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 395
QY 146 GCGCGCGACGACGATGCTGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 205
Db 396 CCGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 452
QY 206 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 265
Db 453 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 512

APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 34443
LENGTH: 1140
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_38459C.1
US-10-437-963-34443

Query Match 14.8%; Score 197; DB 19; Length 1140;
Best Local Similarity 57.4%; Pred. No. 7.1e-43;
Matches 400; Conservative 0; Mismatches 285; Indels 12; Gaps 2;

170 GCGCTCGCCGCGAGCGTCCCGGCGAGGTGCGCGAGAGAGTGCAGGACACTACGAG 229
951 GAGCGGGGTGGCGCTGCTCTCCCGGCAAGACGTGGCCGACGTATGACGACACTACGAC 892
230 GCGCTGTGAGAGAGTGCAGGCGGCGCATGCGCGGCGCGCTCCGCGCTACGCGC 289
891 GACCTCGAGACGAGAGTGTGCTTCATCGAGGCGGCGCTGCTCCCTCCCACTACGCGC 832
290 GGGAGAGAGTCCG 349
831 G-----CCG 781
350 GCGCGACACCG 409
780 GATACACCG 724
410 TGCTCCAGGCGCGAG 469
723 GCG 664
470 CTGCTTCTGCTGGGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
663 TTGTTCTTGAATGGGCTGAG 604
530 TTGCTCATCTCGCGAG 589
603 TTGCTGACGAGCGCGAG 544
590 CTCAACTTCATGAACCGCGAG 649
543 CTCAACTCGCGCGCGAG 484
650 GCGCGCGAGTCAAGTCCG 709
483 CTCCCGCGAG 424
710 AACCG 769
423 CACT 364
770 GCG 829
363 CCTCG 304
830 GCGCGAGAGTGTCCCG 866
303 GCCCAGGTGAAGATCGAGCGCGCGCAATTGCAACGTCG 267

RESULT 10
US-10-437-963-47024
Sequence 47024, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovacic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 47024
LENGTH: 441
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_49833C.1
US-10-437-963-47024

Query Match 14.3%; Score 189.6; DB 19; Length 441;
Best Local Similarity 79.9%; Pred. No. 5.4e-41;
Matches 275; Conservative 0; Mismatches 59; Indels 10; Gaps 4;

1 GTGGAGATCCACCAACCGATGACTCCAGGCGGCGAGAGAGAGAGAGAGAGAGAGAGAG 60
97 GTGGAGATCCACCAACCGATGACTCCAGGCGGCGAGAGAGAGAGAGAGAGAGAGAGAG 156
61 GCG 120
157 GCG 216
121 GCG 172
217 GTGTCACACCG 276
173 CTCGCCGCGAGCG 231
277 TCCCGAGAGGCG 336
232 GCTGATGAGAGAGCGTCCG 290
337 GCTTGTGAGAGAGCGTCCG 396
291 GGAAGAGATCCCG 334
397 GGAAGAGATCCCG 440

RESULT 11
US-10-425-115-183322/c
Sequence 183322, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovacic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 183322
LENGTH: 1793
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_98775C.1
US-10-425-115-183322

Query Match	13.6%;	Score 180.4;	DB 20;	Length 1893;
Best Local Similarity	60.4%;	Pred. No. 2.6e-38;		
Matches 361;	Conservative 0;	Mismatches 216;	Indels 21;	Gaps 3;

OY	198	GCCTGCGCGAAGAGAGTGTGGAGAGCATACAGAGCGCTGTGTGAAGAGCTCCGCGCATTCG	257
Db	495	GGACGCTGTGCGAAGTGTCTCAGCCATTTCAAGACCTTCGAGGTGTGACTTCAGCAGATTCG	554
OY	258	ACGCGGCGCGGTCCTCCGCTCCCGCGCTACGCGCGGAGAGATCTCGCGCGCGCCGACG	317
Db	555	AGACCGGCGAGGTGTGCTTGTCCCGCTTACGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	614
OY	318	GAGCCGAGCCGCGCGCGCGCGCTTCAAGACGCGCGACACCGCGCGACGAGCGCAAG	377
Db	615	GGGACGCTACGCGCCGCTCCGG---GGACTTCAGGCAAGGATCCGCTTCGCGGAG	671
OY	378	GCCTGCGCGCGCGGTACGACCGCGCGCAAGACTGCTCCAAAGCGGAGTGAAGAGCGCA	437
Db	672	GCTCGGGGAGGCGGACCAACGCGCGCAGC-----CGAGCAGAGCGGAGAGA	719
OY	438	AGGCGATCCATYGAACGAGGAGAGACACAGGCTTTCTTGTGGGCGTGGACAAATTTCG	497
Db	720	AGGCGTGTCCATGACGAGAGAGAGACAGGTTTCTCTCTAGGCTTGAAGAAATTACG	779
OY	498	GCAAAGGAGACTGTGCGAGACATCTTCGCGCACTTCGTATCTCGCGACGCGCAACGACAG	557
Db	780	GCAAGGCTGACTGAGAGAACTTTTCGCGCACTTCGTCCAAACCGACGCGCCACCGCAGG	839
OY	558	TGGCGAGCGACCGCGAGAAATCTTCATCCGCTTCAATCTCAATGAACCGCACCGCGCGC	617
Db	840	TGGCGAGCGACCGCGAGAAATCTTCATCAAGCTTCACTTCGCGCGGCAAGAGCAAGAGAGA	899
OY	618	GCTCCAGATCCACGACATCAACAGCTC-----ACCGCGCGGATCAGGTTCGCGCGC	671
Db	900	GCTCCAGATCCACGACATCAACAGCTTCACTTCGAGCGAGACCAAGCGCGCTTCGCGCT	959
OY	672	AGCAGGCGCGCCCGATTCACCGCGCACCGGCGCACCGGCGAACCCCGCGCGCGCGCGCTGG	731
Db	960	CCGAGTCTCTCTTATTCACGAGCGAATGAAAGCACAGCTTCAGCTCTGGCGCGGAGGA	1019
OY	732	GCCGCGCGGCGATGAAGACCAACCAACCAACCCCGCGCGCGCGCGCGCGCTCAT	789
Db	1020	TATGCAAGTCCATTGGCGCGCGCGCAAGCAACGATGCGGAACTCTGCGCTT	1077

Search completed: July 29, 2005, 22:49:37
Job time : 954 secs

QY 490 CAATTGGCAAGGGGAGCTGGGAGCATCTCGCGCAACTTCTCATCTTCGGGAGCC 549
Db 443 GAAATTGGTAAAGAGACTGGCTGAGGAGTATCTTGACTTGTGACTCAAGGACTCC 502
QY 550 AACGAGGTGGCGAGCCAGCGGCGAAGTACTTCTCATCGGCTCAACTGAAACCGCGA 609
Db 503 TACTCAAGTGGCAAGCCATGCGCCGAAAGTATTATTCGGGAGAGTAATGCTGGCCGAG 562
QY 610 CCGCGCGCTCCGACATCCAGCAT 636
Db 563 AAAGAGCGCTCCAGCCTTTTGTACAT 589

RESULT 2
US-09-640-211A-232
; Sequence 232, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Modifications and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640.211A
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 232
; LENGTH: 1358
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-232

Query Match 8.4%; Score 111.8; DB 4; Length 1358;
Best Local Similarity 52.7%; Pred. No. 7.8e-13;
Matches 267; Conservative 0; Mismatches 237; Indels 3; Gaps 1;

QY 133 CGCGGAGCGAGCGCGCGCGCGAGCACTGTTGCGCGCGCTCGCGGAGCGTGGCCG 192
Db 218 CGCGGAGCGCGCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 277
QY 193 GCGGAGTGGCGGAGAGTGGCGAGGCACTACAGGCGCTGTGAGAGAGCTGCGGCG 252
Db 278 CGCGCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 337
QY 253 CATGAGCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 312
Db 338 AACGAGCGGTGATCATGAAAGAGCGGAGCGGAGGAGGCTGTCTCGCGCGCACTA 397
QY 313 CGAGGAGCGGAGCGCGCGCGCGCGCTCCAGAGAGCGCGGAGCGCGGAGCGCG 372
Db 398 CCACT 457
QY 373 CAAGCGCGCGCGCGCGCGGAGAGAGTGTCTTCAAGCGCGGAGAGAG 429
Db 458 CGAGCGGCTACTGTCTCAAGATCCCGCGCGCGCTCTCTCTCTCTCTCTCTCTCT 517
QY 430 GAGGCGGAGGAGTCCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 489
Db 518 GAGGAG 577
QY 490 CAAGTTCGCAAGGGGAGCTGGCGAGCATCTCGCGCAACTTCTCATCTTCGGGAGCC 549
Db 578 GAAATTGGTAAAGAGACTGGCTGAGGAGTATCTTGACTTGTGACTCAAGGACTCC 637
QY 550 AACGAGGTGGCGAGCCAGCGGCGAAGTACTTCTCATCGGCTCAACTGAAACCGCGA 609
Db 638 TACTCAAGTGGCAAGCCATGCGCCGAAAGTATTATTCGGGAGAGTAATGCTGGCCGAG 697
QY 610 CCGCGCGCTCCGACATCCAGCAT 636

Db 698 AAAGAGCGCTCCAGCCTTTTGTACAT 724

RESULT 3
US-09-640-211A-1420
; Sequence 1420, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Modifications and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640.211A
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1420
; LENGTH: 349
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1420

Query Match 8.0%; Score 106.4; DB 4; Length 349;
Best Local Similarity 66.7%; Pred. No. 6.6e-12;
Matches 152; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 420 CGAGAGCGAGAGCGCGAGGAGCATCCATGAGCGGAGAGAGAGAGAGAGAGAGAG 479
Db 59 CGAG 118
QY 480 TGGGCTGTGACAACTTCCGAGAGGAGAGTGGCGAGCATCTCGCGCACTTGTGATCT 539
Db 119 TGGGCTTAAATAATGAGAAAGTGTGATGAGAAACATCTCCAGAACTTGTGATCA 178
QY 540 CGGAGAGCGCAAGCGAGTGGCGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAG 599
Db 179 CGAGAGAGCGAG 238
QY 600 TGAACGCGAGCGCGCGCGCTCCAGAGTCCAGAGCATCAACAGAGTCA 647
Db 239 GTGAAAGATTAAGAGAGGCGAGCATCCAGATTCACACTGTGA 286

RESULT 4
US-09-640-211A-1484
; Sequence 1484, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Modifications and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640.211A
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1484
; LENGTH: 335
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(335)
; OTHER INFORMATION: n = A,T,C or G
US-09-640-211A-1484

Query Match 7.9%; Score 105.6; DB 4; Length 335;


```
Db      99  GCGCCGCTGCGCGAGGTTGAGAGCCATGTCAGACGACGCGCGCCCTCCGCGCCA 158
Qy      203  GCGGAGAGAGTGCAGAGCACTACGAGCGCTGTGAGAGACGTCCGCGCATCGACGC 262
Db      159  CCGCGCGCGGCGCGCGCGGAGTTATCTGTTCGGGGTCAAGGATGATGATGAGACCC 217
Qy      263  GCGCGCTCCCGCTCCCGCTACGCGCGGAGAGTCCCGCGCGCGCGCGCGAGAGCC 322
Db      218  ATGAGGAAGACGCTGAGCTGAAACCTGTGAGATGACGAGCGCCCGAGAGCGCGCC 277
Qy      323  GAGAGCGCGCGCGCGCTGCAAGAGAGCGGAGACCGCGCGCGAGCGCAAGCGCGG 382
Db      278  CCGCGCGCGCGCGCGCTGCGAGAGAGACGCGGAGCGCGCGCTCCGCTACGCGTCC 337
Qy      383  GCGCGCGGATACGACGCGCGCGAGAGTGTCTCAAGCGCGAGAGAGAGCGCGAGGC 442
Db      338  GCGGACGAAGCGGTG-C-GCGACGCGCTGAAAGCGCGCGCGCGCGCGAGCGAGGA 396
Qy      443  ATCCCATGAGACGAGAGAGACACAGGCTGTCTTGTGGGGCTGACAACTTCGCAAG 502
Db      397  GTCCCGTGAACGAGAGATGACACCGGCTGTCTCTGCTCGGCGTACGAAAGTAAAG 456
Qy      503  GGGGACTGGGAGGACATCTGCGCACTGTCTCATCTCGGAGAGCGCAAGCGAGTGGC 562
Db      457  GCGCATTTGAGAGGCACTTTCAGAACTTTGTGAGACTCGAGACCGGACCGAGTCCG 516
Qy      563  AGCCACGCGCAGAGTACTTCTATCCGCTCACTCACTGAAACCG 606
Db      517  AGTATGCGCCAAAATACTTCTCTCGCGGAGCAACCTTAATCG 560
```

```
RESULT 8
US-09-640-211A-1243
; Sequence 1243, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640.211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1243
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1243
```

```
Query Match      7.3%; Score 96.8; DB 4; Length 684;
Best Local Similarity 66.0%; Pred. No. 5.7e-10;
Matches 140; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
```

```
Qy      425  CAGGAGAGCGCAGAGGATCCCATGAGCGAGAGAGAGAGAGAGTGTCTTCTGCGG 484
Db      391  CAGAGAGAGAGAGAGAGTTCATGAGCGAGAGAGAGAGAGTTCGCAATCTTTAATGG 450
Qy      485  CTGAGACAATTCCGCAAGGGGAGTGGCGAGCATCTTCGCGCAACTTCATCTCCGG 544
Db      451  CTAGAGAAAGATGGGAAAGCGGATTGGAGAGCATCTTCAGGAATATGTCACAGAGA 510
Qy      545  ACGGCAACGAGGTGGGAGCGACGCGCAAGAACTTTCATCCGCTCAACTCCATGAC 604
Db      511  ACCCCAAACCAAGTCGAGATGATGCGCAAAATTTCTTCCGCGAGCGCAAGTCTTAAT 570
Qy      605  CGGACGCGCGCGCTCCAGCATCCAGCAT 636
Db      571  AAGAGAGAGCGCGCTCCAGCTCTTGACAT 602
```

```
RESULT 9
US-09-640-211A-1960
; Sequence 1960, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640.211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1960
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1960
```

```
Query Match      7.2%; Score 96.2; DB 4; Length 599;
Best Local Similarity 51.3%; Pred. No. 7.3e-10;
Matches 250; Conservative 0; Mismatches 233; Indels 4; Gaps 1;
```

```
Qy      121  GCGCGCGCGCGCGCGAGAGCGCGCGCGCGAGAGAGTGTGCGCGCGCTCGCGCG 180
Db      117  GACCCGCTCCCTGCGCGCTGCGCGCGAGAGAGATGCTCCCGACGCGCGCGCACCCC 176
Qy      181  GAGCGTCCCGGCGAGAGTGCAGAGAGAGTGCAGAGAGCACTAGAGCGCTGTGGA 240
Db      177  GAGCGTCCCGGCGAGAGAGTGCAGAGAGAGTGCAGAGAGCACTAGAGCGCTGTGGA 232
Qy      241  GAGCGTCCCGGCGAGAGTGCAGAGAGAGTGCAGAGAGCACTAGAGCGCTGTGGA 300
Db      233  GTTCCGCGTCCCGGCGAGAGTGCAGAGAGAGTGCAGAGAGCACTAGAGCGCTGTGGA 292
Qy      301  CCGCGCGCGCGCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 360
Db      293  TCGATACAGAGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 352
Qy      361  GCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db      353  CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 412
Qy      421  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Db      413  TGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 472
Qy      481  GGGGCTGAGCAAGTTCGCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db      473  GCGATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 532
Qy      541  GCGAGCGCAACGAGTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db      533  CCGGAGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 592
Qy      601  GAACCGC 607
Db      593  CAATCCG 599
```

```
RESULT 10
US-09-616-289-48/c
; Sequence 48, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
```


;; TITLE OF INVENTION: Compositions and Methods for the
;; FILE OF INVENTION: Modification of Gene Transcription
;; FILE REFERENCE: 11000.1021CIU
;; CURRENT APPLICATION NUMBER: US/09/640,211A
;; CURRENT FILING DATE: 2000-08-16
;; NUMBER OF SEQ ID NOS: 2368
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 2038
;; LENGTH: 1055
;; TYPE: DNA
;; ORGANISM: Pinus radiata
US-09-640-211A-2038

Query Match 7.0%; Score 92.8; DB 4; Length 1055;
Best Local Similarity 63.4%; Pred. No. 3.8e-09;
Matches 142; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy 413 TCCAGAGCGGAGCAGAGAGGCGGCAAGGCATCCCATGACGAGGAGAGACAGAGCTG 472
Db 608 TCCAGCAATGCACCGCAGAGAGGAGGCGTCCCATGCTCGAGAGAGACAGAGATG 667
Qy 473 TTCTTGCTGGGGCTGACAAATTCGCGAAGGGGAGTGGCGAGCATCTCGCGCACTTC 532
Db 668 TTCTTGATGAGACTGGAGAAAGCTTGGAGAGGGTACTGGAGAGGATATCCAGAAATTT 727
Qy 533 GTCACTTCGCGGAGCGCAAGCGAGTGGCGAGCCAGCGAGAAAGTAATTATCCGCTC 592
Db 728 GTGACGACAGAAACACCCACACAGTAGCCAGCATGCCGAGAGTATTTCTTAAGGCG 787
Qy 593 AACTCCATGAACCGCGACCGCGCTCCAGCATCCACGACAT 636
Db 788 AGCAATCTTAATMAAGAAACGTGATCCAGTCTTTGATAT 831

RESULT 13
US-09-640-211A-1369
; Sequence 1369, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1369
; LENGTH: 328
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1369

Query Match 6.9%; Score 92; DB 4; Length 328;
Best Local Similarity 64.6%; Pred. No. 4.2e-09;
Matches 137; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Qy 425 CAGAGAGCGGAGCGCATCCATGACGAGAGAGACAGAGCTGTTCTTGCTGGGG 484
Db 115 CAAGAGAGAGAGAGAGAGGTTCCATGACAGAGAGAGAGAACTTCTTGTTGGGG 174
Qy 485 CTGAGCAAGTTCGCGAAGGGGAGTGGCGAGCATCTCGCGCACTTCGATCTCCCGG 544
Db 175 CTGAGAGAGCTTGGAGAGGGTGAATTGAGAGGCACTCTAGAGACTATGAGACAGAGA 234
Qy 545 ACGCCAAGCAGAGTGGGAGCGCAGAGAGTAATTCATCCGCTCAACTCCATGAAC 604
Db 235 ACACCGGCGCAGAGTTCAGAGTATGCTCAGAAATATTTCTCCGCGAAGTGAAGCTTCAAC 294
Qy 605 CGGACCGCGCGCTCCAGCATCCACGACAT 636

Db 295 AAGAAAAAGCGGCTCGAGACTCTTTGACAT 326

RESULT 14
US-09-640-211A-1966
; Sequence 1966, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1966
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1966

Query Match 6.9%; Score 91.4; DB 4; Length 427;
Best Local Similarity 61.6%; Pred. No. 5.9e-09;
Matches 146; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

Qy 400 CGGCAAGCTGCTCCAGGCGGAGCAGAGAGCGGCAAGGCATCCCATGACGAGAGA 459
Db 3 CGGCTCCGCTCCTCCATGCGCGCTGAGAGAGAAAGGTATACCATGACGAGAGA 62
Qy 460 AGAGCAGGCTGCTTCTGCTGGGCTGACCAAGTTCGCGAAGGGGACTGGCGAGCAT 519
Db 63 AGAGCTGAGAGGTTTAAATTTGCTCCAGAAATTTGGTAAAGAGACTGGCGAGGAT 122
Qy 520 CTCGCGCACTTCGATCTCGCGAGCGCCAGCAGTGGCGAGCCAGCGAGAGTA 579
Db 123 AGCTCGGACTTTGTGACTGACAGAGACTCTTAAGTGGCAAGCATCCAGAAATA 182
Qy 580 CTTATCCGCTCAATCCATGAACCGGACCGCGCTCCAGCATCCAGACAT 636
Db 183 TTATATCCGAGAGTAATGCTGCGGAGAAAGAGCGCTCCAGCCTTTTATGAT 239

RESULT 15
US-09-640-211A-553
; Sequence 553, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 553
; LENGTH: 883
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-553

Query Match 6.8%; Score 90.4; DB 4; Length 883;
Best Local Similarity 63.2%; Pred. No. 1.1e-08;
Matches 139; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

```

Qy 428 GAGAGCGCGCAAGGCGATCCATGACGAGGAGAGACAGGCTGTCTTGCTGGGCTG 487
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 444 GAGCGAAGAAAGGTGTGTCATGACAGAGAGAGAACATAGATGTTTTGTAGGTTG 503
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 488 GACAAGTTGCGCAAGGGGACTGTGGGAGCATCTGCGCACTTGTCTCTCGGGAG 547
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 504 CAGAGCTTGGCAAGGTGATTGGAGAGAAATAGCACGAAATTTGTCTAACAACGAA 563
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 548 CCAACGACAGTGGCGAGCGCCAGAGAACTTCACTCGGCTCACTCCATGAACCG 607
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 564 CTTACACAGGTAGCCAGCCATGACACAGAAATTTTATTCACAGAGCAATATGACTGA 623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 608 GACCGCCGCGCTCCAGCATCCACGACATCAACAGCGTCA 647
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 624 AAGAAGAGACGTTCCAGTCTGTTGACATGACGCCGCTGA 663
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: July 29, 2005, 22:33:49
 Job time : 266 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Inc.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2005, 17:21:58 ; Search time 5919 Seconds
(without alignments)
10887.893 Million cell updates/sec

Title: US-10-630-636-1

Perfect score: 1330
Sequence: 1 gtgcgagatccaccaccgca.....gacacaaaaaaaaaaaaa 1330

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 2422767955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb_da.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_srs.*
12: gb_sy.*
13: gb_un.*
14: gb_y1.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1330	100.0	1330	AY151042	Oryza sat
2	1181.4	88.8	142680	AP004611	Oryza sat
3	1167.2	87.8	1329	AK068565	Oryza sat
4	851.4	64.0	53000	AP006233	Oryza sat
5	287.8	21.6	187154	OSJN00107	Oryza sat
6	198.4	14.9	1460	AK068623	Oryza sat
7	198.4	14.9	1465	AK106166	Oryza sat
8	197	14.8	1988	AK068138	Oryza sat
9	179	13.5	1502	AK111988	Oryza sat
10	171.6	12.9	1795	AK067964	Oryza sat
11	140.8	10.6	198102	AC145389	Zea mays
12	138.6	10.4	126532	AC130602	Oryza sat
13	135.6	10.4	138467	AC130612	Oryza sat
14	135.4	10.2	945	AY1519528	Arabidops
15	135.4	10.2	1161	AY086906	Arabidops
16	129.2	9.7	897	AY1519532	Arabidops
17	129.2	9.7	928	AY096571	Arabidops
18	128.2	9.7	1593	AY072090	Arabidops
19	128.2	9.6	1030	AK111571	Oryza sat

C	20	125.8	9.5	134982	8	AP003279	Oryza sat
	21	122.6	9.2	1115	8	HVU303354	Hordeum v
	22	120.6	9.1	1104	8	AK068595	Oryza sat
	23	112.8	9.1	1128	8	AK059813	Oryza sat
	24	119.2	9.0	1162	8	AK061500	Oryza sat
	25	119.2	9.0	1490	8	AK103432	Oryza sat
	26	118.6	8.9	867	8	AY1519533	Arabidops
	27	118.6	8.9	898	8	BT005473	Arabidops
	28	118.6	8.9	1640	8	AK118891	Arabidops
	29	116.4	8.8	897	8	AY1519529	Arabidops
	30	116.4	8.8	928	8	BT006122	Arabidops
	31	116.4	8.8	1645	8	AK118135	Arabidops
	32	115.8	8.7	909	8	AT196776	Malus xia
	33	115.2	8.7	63898	2	OSJN00045	Oryza sat
C	34	115.2	8.6	150206	8	OSJN00083	Oryza sat
	35	114.4	8.6	1204	8	AF239956	Hevea bra
	36	113.8	8.6	808	8	LE243339	Lycopersi
	37	112.8	8.5	1208	8	HVU303355	Hordeum v
	38	110.6	8.3	1287	8	AT151044	Oryza sat
	39	110.6	8.3	1357	8	AK101062	Oryza sat
	40	110.2	8.3	1191	8	AB083027	Glycine m
C	41	110	8.3	100800	8	AC105260	Oryza sat
	42	109.2	8.2	142854	8	AC136216	Oryza sat
	43	108.2	8.2	1009	6	AX927169	Sequence
	44	109	8.2	1149	8	AK061427	Oryza sat
	45	109	8.2	3358	8	AK065594	Oryza sat

ALIGNMENTS

RESULT 1
AY151042 1330 bp mRNA linear PLN 10-NOV-2002
LOCUS Oryza sativa (japonica cultivar-group) transcritpion factor MYB51
DEFINITION Oryza sativa (japonica cultivar-group) transcritpion factor MYB51
ACCESSION AY151042
VERSION AY151042
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 1330)
Lu C.A., Ho,T.H., Ho,S.L. and Yu,S.M.
Three Novel MYB Proteins with One DNA Binding Repeat Mediate Sugar
and Hormone Regulation of alpha-Amylase Gene Expression
Plant Cell 14 (8), 1963-1980 (2002)
JOURNAL
MEDLINE
PUBMED
22172034
TITLE
Direct Submission
AUTHORS
Lu,C.A., Ho,T.H., Ho,S.L. and Yu,S.M.
REFERENCE
Submitted (17-SEP-2002) Institute of Molecular Biology, Academia
Sinica, Taipei, Nankang 115, Taiwan
FEATURES
source
CDS
1..1330
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/db_xref="taxon:39947"
20..940
/note="OSMYB51"
/codon_start=1
/product="transcritpion factor MYB51"
/protein_id="AA063152.1"
/db_xref="GI:24850303"
/translation="MTGQAATTTTAAATAATREDKAFENALAAACAPPADGAP
DDWFFALAAVSPARSABEYRBALEVDVAIDGRVLPYAGESAAPDGG
AAASDGGRRPDERGCGGCGGCSABOERKGTWTEERHPLFGDKPG
KGDNRISRNVISRTPTOVAHOKTFRILNSNRDRRSSTIDITSVTAGDOVAAO
QGAPITGHQATGNPAALAGPGCKHHHHPGAGAPPMPWYSAAPMGHVA
AVGTPVVFPGHAPVVPVGVPAAPAKHQ"

RGP. Protein homologues of the coding regions were searched against NCBI Nonredundant protein database with BLASTP.0. ESTs represent the identified cDNA sequences using BLASTN 2.0 with the corresponding DBJ accession no. and RGP clone ID.

A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein.

The orientation of the sequence is from M13 to -21M13 of the BAC clone. This sequence of OJ1005_B10 clone has an overlap with P0516D04 clone (DBJ: AP003276) at the position 1 to 103,355 of 5' end. The nucleotide sequence of this BAC clone was generated by combining Monsanto and RGP-Japan sequencing data. Detailed information on overlap and assembly quality together with annotation of this entry is available at

<http://rgp.dna.affrc.go.jp/Genomeseq.html>.

FEATURES
source

```
1. 142680
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clone="OJ1005_B10"
join(3399..3512,4006..4227,4265..4328,4613..4792,
6611..6996)
/gene="OJ1005_B10.1"
join(3399..3512,4006..4227,4265..4328,4613..4792,
6611..6996)
/gene="OJ1005_B10.1"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC05637.1"
/db_xref="GI:21902090"
/translation="MGHPCLRTGTARPDSSWMAVPREVOPIGRHGLARSAYKPPRG
GGEGERRRPAPSLGREGREKREEMVTVPDCCSTDCNLIISFPOLPHLSSSPRVA
GGPWPBPRAAEHTGCLKPEVGGSAAGYFEMACYSRGRHGVGRSLPNCR
EPRLATLPCHITAAFPDFGTALDRNKFTYCPSESGAAGASTASREERTSSM
VDTEVGALPVRERRRIWPPRPVADLASRPGGSGGLPEARRRIIPRABELL
SRPRRAASHSVPPPELAAAFNASHVEGVRGREGCEFLQ"
join(9727..9823,11568..11674)
/gene="OJ1005_B10.2"
join(9727..9823,11568..11674)
/gene="OJ1005_B10.2"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC05638.1"
/db_xref="GI:21902091"
/translation="MAEAGCTRASQOEBAQIENGSGKQLPEDTTCVAAAGAAPF
SMSNSVALLMNGSGGSSLPFHE"
complement(13538..14391)
/gene="OJ1005_B10.3"
complement(13538..14391)
/gene="OJ1005_B10.3"
/note="gag-pol polyprotein
probably inactive due to frameshift(s) in CDS
probably inactive due to stop codon(s) in CDS
pseudogene"
pseudogene
join(16900..17014,17370..17504,18435..18718)
/gene="OJ1005_B10.4"
join(16900..17014,17370..17504,18435..18718)
/gene="OJ1005_B10.4"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC05639.1"
/db_xref="GI:21902092"
/translation="WNTTSKVHGDHSGSELVLLGHPATBELQLIQPILPTGSGSIR
KGRVAHPNSIMLTASVTEKALQVQHLFISCHRCQDQDGGSGEGRRRAGQ
PEKAVTAGDGERHGERLAEVALLAHAGAGGEGRETRREGEDLLIKHRAVQERT
```

```
ATTWTLRAKGRGSGDN"
join(22941..24398,24894..25130)
/gene="OJ1005_B10.5"
join(22941..24398,24894..25130)
/gene="OJ1005_B10.5"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC05640.1"
/db_xref="GI:21902093"
/translation="WNTFQTLTMMEAALVACIFISPRRLKQDVYDQKVKQDHK
NIRSLNFTGLAHSMSYITLVAFQPLVAPLRYVHSSYKHKDMQDAVETAK
KLMAFIRYNVEALEMNLITGKELARSDIDDLQGRILIDYITIRWEKHLRSGVL
DIYIEIKYKVEALISIRASTETLERVVMNLNKSLEETRGHAISVLEAPVLL
FENIPAMPOLASLITGERINGEPHDINSEVLCYGVKILERLVENDNRSVADA
NVLSKIVELNIPDDCKVPVDETERNSOEIYASLNVLHKLVTSTGEGALRST
ISKSEISMNIRKILYOHDNKSSLSVAHVYLSCLMHTARFVIGSSCOIVRLVS
SVLRPMDIVDGNSTVADSAQALVYLSTGSEYRATITLDKINLEELVGMFLDAS
MEORIMIAHLKHLRTYSGPGYNQLKVIDGSPPKDKELKAEKGLDSPTGLSSQ
ICHSLOAIDPDKALQSANLTMDTVAODTERVOISYHFPDDKKSNTQSDMAQK"
complement(join(26516..26605,26825..26971))
/gene="OJ1005_B10.6"
complement(join(26516..26605,26825..26971))
/gene="OJ1005_B10.6"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC05641.1"
/db_xref="GI:21902094"
/translation="WENITSRVIFKVLGGDRVVISYKIQCLRVLLGDDDDPYKX
DPQGRILVPTVNTLIRYISITRQLTPQY"
complement(join(28040..28278,28339..28375))
/gene="OJ1005_B10.7"
complement(join(28040..28278,28339..28375))
/gene="OJ1005_B10.7"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC05642.1"
/db_xref="GI:21902095"
/translation="WGAADPAETAAHPAEGRGVFSRRPGORPEGVSAAVRRHG
GGSGVGRSGRIGRRRTAATAVAFVDFDSDPATRATDYG"
join(30138..30218,31910..32079,32238..32400)
/gene="OJ1005_B10.8"
join(30138..30218,31910..32079,32238..32400)
/gene="OJ1005_B10.8"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC05643.1"
/db_xref="GI:21902096"
/translation="MQLNPKERTRRNKKAKGWCABIVLNKAATTIHDIADGVKGG
GGGCTRLDPLMRSTETETAIVAPERSHADGVDSGGCCCRADNSGDGLTAVAPEAFTE
VVDGGSYGGGRIMWRERKGRKREGGEDETDS"
33576..33806
/gene="OJ1005_B10.9"
33576..33806
/gene="OJ1005_B10.9"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC05644.1"
/db_xref="GI:21902097"
/translation="WRRRVYQKSTAAAGADAABAATGEDDEDDGDEGVEDDED
DGGCGHGEAERGEDAAAGAEVAERGGG"
join(35782..36026,36318..36486)
/gene="OJ1005_B10.10"
join(35782..36026,36318..36486)
/gene="OJ1005_B10.10"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC05645.1"
/db_xref="GI:21902098"
/translation="MAVPPKDAAPDOPDNDERLPEKLTMAVRIAPRYVMNC
GALPANAATVVLGAYTLLIKQDPFWYTTIVFEATICPGKFTLTSRLIADTFLDSD
NICVRCRYRRSSFTAGTDLKLTGMRHOSHREL"
complement(join(39042..39337,40177..40346,40543..40590,
41257..41417,42100..42248))
```

```

Query Match      88.8%; Score 1181.4; DB 8; Length 142680;
Best Local Similarity 91.2%; Pred. No. 2e-116;
Matches 1316; Conservative 0; Mismatches 6; Indels 121; Gaps 1;

/ gene="OJ1005_B10.11"
/ complement{join(39042..39137,40177..40346,40543..40590,
41257..41417,42100..42248)}
/ gene="OJ1005_B10.11"
/ note="hypothetical protein"
/ codon_start=1
/ protein_id="BAC05646.1"
/ db_xref="GI:2190209"
/ translation="MARTGAAVAHAHTDGLATFNMTGKYELGCTSSDPEFPAAH

OY      1  GTGCCAGATTCACCAACCCGATGACTCTTCCAGGCGGGGAGAGACGACGACCAACCGCGCGCGC 60
DB      138939  GTGCCAGATTCACCAACCCGATGACTCTTCCAGGCGGGGAGAGACGACGACCAACCGCGCGCGC 138998
OY      61  GCGCGGCGGGGTGGACCAAGGAGAGACGACAAAGGGGTTGAGAAAGCGCTGCGGGCTTGCGC 120
DB      138999  GCGCGGCGGGGTGGACCAAGGAGAGACGACAAAGGGGTTGAGAAAGCGCTGCGGGGAGAGGTC 139058
OY      121  GCGCGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 180
DB      139059  GCGCGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 139118
OY      181  GAGCGTGTCCCGGGGCGAGGTTCGCGGAGAGAGGTTCGCGAGGACCTACGAGGCGCTGTGTGA 240
DB      139119  GAGCGTGTCCCGGGGCGAGGTTCGCGGAGAGAGGTTCGCGAGGACCTACGAGGCGCTGTGTGA 139178
OY      241  GGAAGCTGCGGGGCGCATGAGCGCGGGGCGCGCTCCGCGCTACGCGCGGAGAGAGTGC 300
DB      139179  GGAAGCTGCGGGGCGCATGAGCGCGGGGCGCGCTCCGCGCTACGCGCGGAGAGAGTGC 139238
OY      301  CGCGGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCCAGAGCGCGAGACCG 360
DB      139239  CGCGGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCCAGAGCGCGAGACCG 139298
OY      361  GCGCGAGCGACCGGAGGCGCGCGCGCGGGTACGACGCGCGGACGAGAGCTGTTCAAAGGC 420
DB      139299  GCGCGAGCGACCGGAGGCGCGCGCGCGGGGTACGACGCGCGGAGAGAGCTGTTCAAAGGC 139358
OY      421  GGAGCGAGGAGGCGGAGGCGCATCTCCATGAGCGAGGAGAGACACAGTACGCGCGATT 485
DB      139359  GGAGCGAGGAGGCGGAGGCGCATCTCCATGAGCGAGGAGAGACACAGTACGCGCGATT 139418
OY      486  ----- 485
DB      139419  CGGGAATATAGACGAGCAGCTGTTGTTGCTTCTTCTTCGGAACAGCTCGATTGC 139478
OY      486  -----CAGGCTGTTCTTGC 479
DB      139479  GAATCCATGCGCGGAGCTGATCTGATTGCGATTGAGTTGCGCGGAGGCTGTTCTTGC 139538
OY      480  TGGGGGTGACAAATTTCGCGAAGGGGAGCTGGGGAGAGCATTCGCGCAATTGTGTATCT 539
DB      139539  TGGGGGTGACAAATTTCGCGAAGGGGAGCTGGGGAGAGCATTCGCGCAATTGTGTATCT 139598
OY      540  CGCGGACGCGCAACGACGAGTGGCAGAGCGACGCGGAGAAATGTTACTTTCATTCGCGTCAACTCCA 599
DB      139599  CGCGGACGCGCAACGACGAGTGGCAGAGCGACGCGGAGAAATGTTACTTTCATTCGCGTCAACTCCA 139658
OY      600  TGAACCGGAGCGCGCGCGCTTCAGAGTCCAGACATCAACGACATCAACGAGCTGACCGCGCGATC 659
DB      139659  TGAACCGGAGCGCGCGCGCTTCAGAGTCCAGACATCAACGACATCAACGAGCTGACCGCGCGATC 139718
OY      660  AGGTGCGCGCGGACGAGGAGGCGCGCGATCAACGCGCGACAGAGGCGAACC CGCGCG 719
DB      139719  AGGTGCGCGCGGACGAGGAGGCGCGCGATCAACGCGCGACAGAGGCGAACC CGCGCG 139778
OY      720  CGCGCGCGCGCGCGCGCGCGGAGATGAAGACCAACCAACCAACCAACCGCGGCGCGCGCG 779
DB      139779  CGCGCGCGCGCGCGCGCGCGGAGATGAAGACCAACCAACCAACCAACCGCGGCGCGCGCG 139838

```

Oy		780	CGCCGCCCATGCCCCATGTAAAGACGGCGGCCCATATGGGCACCCTCGTGGCGGCACATGG	839
Db		139839	CGCGGCCCATGCCCATTGAACAGCGCGGCCCATATGGGCACCTCCGTGCGCGGCACATGG	1398988
Oy		840	TGCCCGCCCGCGTCGGCAAGCGCGTGTTGTTCCCGCGGGCCACGCGCGTAAGTGTGC	899
Db		139899	TGCCCGCGCGCGTGGCAGCGCGGTGTGTTCCCGCGGGCCACGCGCGGTAGTGTGTGC	1399588
Oy		900	CCGTGGCTACCCCGCGCGCTCCGCGCAAAGATGACCAATGACGCGCGCATGACAGCATG	959
Db		139959	CCGTGGCTACCCCGCGCGCACCGCGCAAAGATGACCAATGACGCGCGCATGACAGCATG	140018
Oy		960	AGCAGCATTTCTTCTCTCCCTTCTTGATGTCAATCTGATTGTCTTGTGTGTG	1019
Db		140019	AGCAGCATTTCTTCTCTCTCTTCTTGTATGTGAATCTGATTGTCTTGTGTGTG	140078
Oy		1020	CCGGCTCATGCTCCCTGATCATCTTGTCTTTCACAATCTCATAATGTAACATACAT	1079
Db		140079	CCGGCTCATGCTCCCTGATCATCTTGTCTTTCACAATCTCATAATGTAACATACAT	140138
Oy		1080	AGATCAGATGCCAAGAAGTGCAGGGATTGGGATTAAAGCGAATAATGAATTTTGC	1139
Db		140139	AGATCAGATGCCAAGAAGTGCAGGGATTGGGATTAAAGCGAATAATGAATTTTGC	140198
Oy		1140	TGACTGTTTTGAAGATCATCATCGTACACCGCGTGAATAGCTTAGCTCCAATGTGGATG	1199
Db		140199	TGACTGTTTTGAAGATCATCATCGTACACCGCGTGAATAGCTTAGCTCCAATGTGGATG	140258
Oy		1200	TAAATTAGCAGCGCGCTTCGTACGTGTGGCGCCCATGATGATCTTGCAGGGGTTTCAA	1259
Db		140259	TAAATTAGCAGCGCGCTTCGTACGTGTGGCGCCCATGATGATCTTGCAGGGGTTTCAA	140318
Oy		1260	TTAGGAGATTGATTTTCATTTTGCTGATGTAATTTGCCAACGTCTCATTTGACAAAAA	1319
Db		140319	TTAGGAGATTGATTTTCATTTTGCTGATGTAATTTTGCTCATTTGACAAAAA	140378
Oy		1320 AAA 1322		
Db		140379 ATA 140381		
RESULT 3 AK068565 LOCUS DEFINITION ORYZA sativa [japonica cultivar-group] cDNA clone:J013154612, full insert sequence. ACCESSION VERSION KEYWORDS SOURCE ORGANISM REFERENCE AUTHORS		AK068565 AK068565 AK068565.1 GI:32978582 FLI_CDNA; CAP trapper. Oryza sativa [japonica cultivar-group] Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehharitoidae; Oryzaceae; Oryza.	1329 bp mRNA linear cDNA clone:J013154612, full insert sequence. PIL-CDNA; CAP trapper. Oryza sativa (japonica cultivar-group) Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehharitoidae; Oryzaceae; Oryza.	PLN 24-JUL-2003
The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team; Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yarakai,J., Ishikawa,M., Yanada,H., Ooka,H., Hotta,I., Kojima,K., Nemiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Onheuki,K., Shieniki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group.; Okono,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kuratsaki,T., Kodama,T., Maesuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., NIKURA,J., Ikeda,R., Ishibiki,Y., Kawamata,M., Yoshimura,A., Mura,Y., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kasawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Salto,R., Sasaki,D., Sato,K., Shibata,K., Shingawa,A., Shiraki,T., Yoshino,M., and Hayashizaki,Y.				

TITLE Collection, mapping, and annotation of over 28,000 cDNA clones from
Japonica rice
JOURNAL Science 301 (5631), 376-379 (2003)
MEDLINE 22752273
PUBMED 12869764
REFERENCE 2 (bases 1 to 1329)
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirao, T.,
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Kuroski, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
Miyano, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
Nanuki, T., Narikawa, R., Nishikawa, Y., Nishikawa, Y., Nomura, K.,
Nunasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
Osato, N., Ota, Y., Otomo, Y., Ryu, R., Satoh, H., Sakai, C., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
Tagawa, A., Takanashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yanagi, W.,
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
Yoshimura, A.
Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agricultural Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp)
Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28k full-length cDNA clones from Japonica
rice.
URL: <http://cdna01.dna.affrc.go.jp/cdna/>
NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Nanuki, T.,
Ohneda, E., Yanagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and
Yamamoto, M.
FAS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
Kodama, T., Kuroski, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,
Miyano, K., Narikawa, R., Nishikawa, Y., Oka, M., Ryu, R., Sugano, S.,
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
Hirao, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Nunasaki, R., Ohno, M., Osato, N.,
Ota, Y., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takanashi, F.,
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A. and Hayashizaki, Y.
FEATURES
source location/Qualifiers
1..1329
/organism="Oryza sativa (Japonica cultivar-group)"
/mol_type="mRNA"
/cultiivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J013154G12"

ORIGIN
Query Match 87.8%; Score 1167.2; DB 8; Length 1329;
Beet Local Similarity 94.7%; Pred. No. 2.1e-114;
Matches 1245; Conservative 0; Mismatches 3; Indels 66; Gaps 1;

QY 1 GTGCGAGATCCACCAACCGATGACCTCCAGGCGGCGACGACGACCAAGCGGCGGC 60
DB 82 GTGCGAGATCCACCAACCGATGACCTCCAGGCGGCGACGACGACCAAGCGGCGGC 141
QY 61 GCGCGCGCGGTGACACGAGGAGACGACGAGCGCTTCAGAAACCGCGCTTGCGGC 120
DB 142 GCGCGCGCGGTGACACGAGGAGACGACGAGCGCTTCAGAAACCGCGCTTGCGGC 183
QY 121 GCGCGCGCGCGCGGACGAGGCGGCGCGGACGAGCTGTTCCGCGCGCTGCGGC 180
DB 184 -----CGCGCTGCGGC 195
QY 181 GAGCGTCCCGGCGGCGAGTTCGCGGAGAGTTCGAGGACCTACGAGCGCTGTTGCA 240
DB 196 GAGCGTCCCGGCGGCGAGTTCGCGGAGAGTTCGAGGACCTACGAGCGCTGTTGCA 255
QY 241 GAGCGTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
DB 256 GAGCGTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 315
QY 301 GCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
DB 316 GCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 375
QY 361 GCGCGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
DB 376 GCGCGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 435
QY 421 GAGCGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
DB 436 GAGCGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 495
QY 481 GGGCGTGAAGAGTTCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
DB 496 GGGCGTGAAGAGTTCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 555
QY 541 GCGGACCCCAAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
DB 556 GCGGACCCCAAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 615
QY 601 GAAACCGGACCGCGCGCGCTTCAGACATCCAGACATCCAGCGCTACCGCGCGATCA 660
DB 616 GAAACCGGACCGCGCGCGCTTCAGACATCCAGACATCCAGCGCTACCGCGCGATCA 675
QY 661 GGTGCGCGCGGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
DB 676 GGTGCGCGCGGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 735
QY 721 GCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
DB 736 GCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 795
QY 781 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 840
DB 796 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 855
QY 841 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 900
DB 856 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 915
QY 901 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 960
DB 916 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 975
QY 961 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1020
DB 976 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1035
QY 1021 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1080
DB 1036 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1095

QY 1081 GATGAGTCCAGAGAGTGCGAGGATTGGGATTAAAGCGATTAAGTATTTGGCT 1140
DB 1096 GATCAGATGCGCAGAGAGTGCGAGGATTGGGATTAAAGCGATTAAGTATTTGGCT 1155
QY 1141 GACTGTTGGCAAGGATCATCAGGTACACCGCGTGAAGGCTTAGCTCCAAATGTTGATGT 1200
DB 1156 GACTGTTGGCAAGGATCATCAGGTACACCGCGTGAAGGCTTAGCTCCAAATGTTGATGT 1215
QY 1201 AATTAGCAGCGGCGCTTCGTAACGTGTGCGCGCGATGATGATCTTCGAGGGTTGCAAT 1260
DB 1216 AATTAGCAGCGGCGCTTCGTAACGTGTGCGCGCGATGATGATCTTCGAGGGTTGCAAT 1275
QY 1261 TAGGAGTGAATTTCCATTTTGGTCTGATGTAATTTGCCAATCTGTCTCATTTGGACC 1314
DB 1276 TAGGAGTGAATTTCCATTTTGGTCTGATGTAATTTGCCAATCTGTCTCATTTGGACC 1329

RESULT 4
AP006233 53000 bp DNA linear PLN 31-AUG-2004
LOCUS Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
DEFINITION BAC clone: B1249E06, complete sequence.
ACCESSION AP006233.2 GI:29824107
VERSION HTG.
KEYWORDS Oryza sativa (japonica cultivar-group)
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehmeritoidae; Oryzaceae; Oryza.

REFERENCE
AUTHORS Sasaki, T., Matsumoto, T. and Katayose, Y.
TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC
clone: B1249E06
JOURNAL Published Only in Database (2003)
AUTHORS 2 (bases 1 to 53000)
TITLE Direct Submission
SUBMITTED (19-FEB-2003) Takuji Sasaki, National Institute of
Agricultural Sciences, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: sasaki@nias.affrc.go.jp; URL: http://rgp.dna.affrc.go.jp/
Tel: 81-298-38-7441, Fax: 81-298-38-7468)
COMMENT On Apr 11, 2003 this sequence version replaced gi:28460672.
The orientation of the sequence is from -21M13 to M13rev of the BAC
clone.

FEATURES
source location/Qualifiers
1..53000
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clone="B1249E06"

ORIGIN
Query Match 64.0%; Score 851.4; DB 8; Length 53000;
Best Local Similarity 98.7%; Pred. No. 1.7e-81;
Matches 858; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

DB 230 CATCACGAGGTACCGCCGCGATGATGATGCGCGCAGACAGGGCCCGCATCACCG 289
QY 694 CCACAGGCGCACCGGCAACCCCGCGCGCGGCTGAGGCCCGCGGCAATGAAGCAACA 753
DB 290 CACACAGGCGCACCGGCAACCCCGCGCGCGGCTGAGGCCCGCGGCAATGAAGCAACA 349
QY 754 CCAACACCAACCCCGCGCGCGCGCGCGCGCCCATATGATATACAGCGCGCCCAT 813
DB 350 CCAACACCAACCCCGCGCGCGCGCGCGCGCCCATATGATATACAGCGCGCCCAT 409
QY 814 GGGCGACCCCGTGGCGCGCGCACATGAGCCCGCGCGCGCGCGCGCGCGCGCGCGCG 873
DB 410 GGGCGACCCCGTGGCGCGCGCACATGAGCCCGCGCGCGCGCGCGCGCGCGCGCGCG 469
QY 874 GCCGCGCACCGCGCGCGCGCACATGAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 933
DB 470 GCCGCGCACCGCGCGCGCGCACATGAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 529
QY 934 CCAATGACCGCGCATGAGCGGACATGAGAGCAATTTCTTCCTCTCTTCTTATGATGCA 993
DB 530 CCAATGACCGCGCATGAGCGGACATGAGAGCAATTTCTTCCTCTCTTCTTATGATGCA 589
QY 994 ATCTGATTTGTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1053
DB 590 ATCTGATTTGTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 649
QY 1054 ACAATCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1113
DB 650 ACAATCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 709
QY 1114 AAAGCGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1173
DB 710 AAAGCGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 769
QY 1174 TGAAGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 1233
DB 770 TGAAGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 829
QY 1234 GATCGATGATCTTTCGAGGCGGTTGCAATTAAGGATTTGATTTGATTTGATTTGATTTGATTT 1293
DB 830 GATCGATGATCTTTCGAGGCGGTTGCAATTAAGGATTTGATTTGATTTGATTTGATTTGATTT 889
QY 1294 TGGCAACTGTCTGATTTGACCAACCAAAAAA 1322
DB 890 TGGCAACTGTCTGATTTGACCAACCAACATA 918

RESULT 5
OSJN00107/c 187154 bp DNA linear PLN 10-FEB-2004
LOCUS Oryza sativa genomic DNA, chromosome 4, BAC clone: OSJNB0062B06,
DEFINITION complete sequence.
ACCESSION AL606729 GI:32488976
VERSION AL606729.3
KEYWORDS HTG.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehmeritoidae; Oryzaceae; Oryza.

REFERENCE
AUTHORS Peng, Q., Zhang, Y., Hao, P., Wang, S., Fu, G., Huang, Y., Li, Y., Zhu, J.,
Liu, Y., Hu, X., Jia, P., Zhang, Y., Zhao, Q., Ying, K., Yu, S., Tang, Y.,
Wang, Q., Zhang, L., Lu, Y., Mu, J., Lu, Y., Zhang, L., S., Yu, Z., Fan, D.,
Liu, X., Lu, T., Li, C., Wu, Y., Sun, T., Lei, H., Li, T., Hu, H., Guan, J.,
Wu, M., Zhang, R., Zhou, B., Chen, Z., Chen, Z., Jin, Z., Wang, R.,
Yan, H., Cai, Z., Ren, S., Lv, G., Gu, W., Zhu, G., Yu, Y., Jia, J.,
Zhang, Y., Chen, J., Kang, H., Chen, X., Shao, C., Sun, Y., Hu, Q.,
Zhang, X., Zhang, W., Ding, L., Chen, C., Sheng, H., Gu, J., Chen, S.,
Mi, L., Zhu, F., Chen, W., Lan, L., Lai, Y., Cheng, Z., Gu, M., Jiang, J.,
Li, J., Hong, G., Xue, Y., and Han, B.
TITLE Sequence and analysis of rice chromosome 4

JOURNAL Nature 420 (6913), 316-320 (2002)
MEDLINE 22373737
PUBMED 12447439
REFERENCE 2
AUTHORS Han, B., Feng, Q., Huang, Y.C., Li, Y., Zhu, J.J., Zhao, Q., Hu, X., Liu, Y.L., Mu, J., Yu, Z., Chen, L., Fan, D.L., Weng, Q.J., Zhang, L., Liu, Q., Yu, S.L., Liu, X.H., Lu, T.T., Zhang, Y.J., Lu, Y., Li, C., Li, T., Zhang, Y., Hu, H., Jia, P.X., Qian, Y.M., Yang, K., Zhou, B., Chen, Z.H., Hao, P., Zhang, L., Wu, M., Zhang, R.Q., Guan, J.P., Fu, G., Wang, S.X., Ren, S.X., Lv, G., Lin, W., Gu, W.Q., Zhu, G.F., Tu, Y.F., Jia, J., Yin, H.F., Zhang, Y., Cai, Z., Chen, J., Kang, H., Chen, X.Y., Shao, C.Y., Sun, Y., Hu, Q.P., Zhang, X.L., Zhang, W., Wang, L.J., Ding, C.W., Sheng, H.H., Gu, J.L., Chen, S.T., Ni, L., Zhu, F.H. and Hong, G.F.
TITLE Direct Submision
JOURNAL Submitted (08-SEP-2001) Han Bin, National Center for Gene Research, Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233, CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests: bhan@ncgr.ac.cn
REMARK Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC clone: OSJNB0062B06.
COMMENT On Jul 9, 2003 this sequence version replaced gi:21912495.
Web site: http://www.ncgr.ac.cn
----- Summary Statistics -----
Assembly program: phrap

This is a complete sequence.
Genes were identified by a combination of several methods: gene prediction programs including Fgenesh (http://www.softberry.com/), GENSCAN (http://CCR-081.mt.edu/GENSCAN.html), GenemarkHM (http://genemark.biology.gatech.edu/genemark/), tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/), searches of the complete sequence against NCBI non-redundant protein database (nr) (ftp://ncbi.nlm.nih.gov/blast/db) and the EST database at NCGR.
Location/Qualifiers
1. 187154
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:39947"
/chromosome="4"
/clone="OSJNB0062B06"
/clone_1ib="CUGI-OSJNB"
3932..5799
/gene="OSJNB0062B06.1"
join(3932..4688,5168..5282,5395..5761,5847..6161,6241..6561,6647..6799)
/gene="OSJNB0062B06.1"
/codon_start=1
/protein_id="CAE04043.2"
/db_xref="GI:38345960"
/translation="MTTTLAEVYVMVHGVGLPLVGVIFLHGTCTRYFDFRQAILPSMPPNSILFTGMOKLEBLRKKAMKRAALVQGOORFELICODKAGRIYRKYKCEVYKANGTCHGSCAKPKLHRPCTHYIAAAGEGIDPAVSYGYSKGIAYHTWS EITGVGIAGERTNDEVLTNIPDSKIRGAKGAPRTBRIRNDMDSDAGRYKSCDC ERGHTYKCPDKEKRNALBAGLSSANDGAPRTDTGTGTSVVGGRSSTPSPYC GTGSDVDVPGTVDVLRAGLLPMCRIVEAAGDVDPARRVTWDSLLAALVDRRPF THTPLPCGEVAPLTLDISYLLGLADAVAPVTGVWDLDLFAFALVQAPHLPLEPLAHNTGTPTKMLQFTVEQLAGADEYSRCLEAYLLMFGVMFCGGHGH YKSYRLSMAARRGVEDPMWSDPCVYLLPMCPRLACGRCOLMASRPLDGRH PRCQNVDRRTTACAYIEWVDTEPNPLDITTCLOEGRKWFASSTQYIQRKAYE RQCEQOSDNRVLITLALPPEAKPRCRGCDRCQVRSIKPTTLGRFFVCALITDDDF MEPPRCQYREMDITRRVLTTPSHVQLEPQYKVTARFERGSSRRG"
gene
CDS
complement(join(10043..10336,10864..11112,13427..13562,16168..16386,17655..18033,18577..18817))
/gene="OSJNB0062B06.2"
/codon_start=1
/protein_id="CAE04044.2"
/db_xref="GI:38345961"

/db_xref="UniProt/TREMBL:Q7X7F4"
/translation="MERADKLKQDVRILFETRNHTAKKMLVDAVORIGLIDHLPDEI NNVIDKSEFTSSSLBEVALKDFRLRREHIIWSPVFEKEDGFIITIDEPRA ALISLVNAHLVHDEPDLBEAMSFAHHHSDGSRILKALNOLNIRALPLPRT YKRYEMLHLYEQEEECIVYLDLDAKLEBNLLOHNLKELKAFSOMKDWXCYVEL SHVRDRAVESTLWSYALFEENLTLTLMILAKIIVFVLMDDTYDDHATIECRKLE AIQWDBSASISLPEKRYKFAQLQNFRETEAVGSDKRVYCMKPEFNLSTYLL OEPMHLQNYPAPEKREVALSTSTSYPLCVTAVGQDAVTESFELTVRSAAVI ACATIEFMDIAPFKSGRKNKGDAANTVECYENENKTSVALDKIESMSEWRTL NQVCDHHDQPPVQVRLNLNLSVPFFYDKKDAVYTSRYIQEIVGDLFVNVPVPI"
26861..27649
/gene="OSJNB0062B06.3"
26861..27649
/gene="OSJNB0062B06.3"
/codon_start=1
/protein_id="CAE04045.2"
/db_xref="GI:38345962"
/translation="MVFPLNDLFLPRADLELYLTDMFSTQNEIDSLITLIPADKVPDGFNGFPMKWHIIIAHFAEATDLOVNLNYSYETLQTLAQAPEYIHOCQKRIIILDKP LTKLMADRLQVLEVVGENQYEFIKRTIOECLAQAPEYIHOCQKRIIILDKP EKADPTIEHTLILSVMENMGPRKMLWMQWAFSSASTILLNGVSGISFRCXGVRO GDPISLPLFVGAELQRIYNRAFNILGKLPETYIGLRLVLDQK"
complement(36241..36855)
/gene="OSJNB0062B06.4"
complement(36241..36855)
/gene="OSJNB0062B06.4"
/codon_start=1
/protein_id="CAE04046.2"
/db_xref="GI:38345963"
/translation="MPTTVAQRGELARNGRDEESTRTNGKDDGDTGCGARP RVTLPLRCRRRRRRRAHGHGCGSGCGRSDGCGVGLRTNLRAGCGRRR GAGGDDDVGRSCAATGRCLEAAGMBRGMBRGCHVTLRRLRKGRNRGHFI GAGRPATGRAGCGNGRSMKLSGIEBEGGIEPPLINLE"
complement(45605..50168)
/gene="OSJNB0062B06.5"
complement(45605..50168)
/gene="OSJNB0062B06.5"
complement(join(45605..45901,46925..47173,47477..47612,48065..48283,49052..49430,49925..50168))
/gene="OSJNB0062B06.5"
/codon_start=1
/protein_id="CAE04047.2"
/db_xref="GI:38345964"
/translation="MVRRAEKLKQDVRITMEPTCSTEGRLQVLAOLHIGLIDHLEKE IEISLSEINSEPISSLDHVALFRLRLRGFRVSPVFRKFKDDDRPFGITNDP RGLSLVNAHLTHDEPELEASLFTQHLASGSDVDLNPHLIDQINRLDVLPRTYRMERTLCVMEPEYRQEGHPILEILAMDLFNLOHVLKELKAISEWMDLYGML LSYLRDVSYSYVSVYFVEEDSALAMITFKIIAFTIIMDDTYDSAYTIOECRLN EAIORMBESATAPLEPYIKFYSALTKTFKEFEIIVBEDDQGYRIDHTKAPONSAY LOELEMYSQNKPSPEEQVALSYTSYPLICVETVGRGALINAEAPMANDIGAK IACKITRFNDINAFKQGRKRGDVAVSTECYMEKVTSEGAFTKIDIMTEBERK TTNALCEHRELAFAVOQVNLAICAFTEYGRKDAVTFPSHLQETVESLFRPVSI"
complement(64764..71118)
/gene="OSJNB0062B06.6"
complement(join(64764..64805,64958..65940,66221..67084,67774..67858,68435..68536,68753..69149,69306..69546,69909..70953,71089..71118))
/gene="OSJNB0062B06.6"
/codon_start=1
/protein_id="CAE04048.2"
/db_xref="GI:38345965"
/translation="MTVLRVNPKAPAPAPADVNSNSAVTTPKQGLSKTNGVSVH KRLQPVADSLPPEGFSKVSASSVLSASVRI SVSHPEEASVLSBSSSVSEKAK YMRKQSSRSSTSLERTRLRRAHFLHVMGRCFRCLESGRVIDPLXLSVSVSEGL DMKSSAPRRDLALNPOMLPLVPLCFQOLSPPHLSLVASALRQASRRRPPMSS MCAGOTPLLRDLPLGRSSATMBDLICVPLMRGMLTTRREGOMRLADMTFRAPRLH RDOVQNKYVEDSPSYLSRADAHLICQRLDCHGRKILMEVNVHILAIAROBAAHHVYR YRARCMEGIMPMAKTEAVAKLIGRPVTCGCGKKILMEVNVHILAIAROBAAHHVYR VSRAGRAVTVGSMIKASTTRTTVSAATKTPLEPAALNTPGAELRKPLVPPQVR VOHSSSPSHRLKPTQATNLSPCNVKRPQSSKSPRGAGALLVPPSPPELFFLHQ DNCNGQHDLDPMLEAVINYS PDGCPGEGSSPYVAAPTAEGALVAGSDVNASADLLA

SLATPRFKQPTPSRLMADLYEKGCKKIKIVAGKNNFFLPADGTBGGIILAA
DDNYRSLQSOETNLSPEEALARSIAKATLGAALANRIKIRORSULTYIKEDANT
KLPHIKANARRKNFIHYLOTSTGFALSHOKEDELDNPFSDRGKOSRELSMWS
LGMSPDLSDOEEATIMEBELATIKSPSKAPKPPDPSFSAFYKCMDLIKHDLFNAY
RSFPLKNONOLDLNTCHICLPPKADASAPEHRPISLIHSEFKIISKLANLAPR
MAELVTANQSAFIRKRAIHDFNLYVQULYKFRKRNPTLMKHIDISADAYNAYM
MEVLSVJDAOSIFNFKOLNLFKASGLQNNVHTQULPACBSLNVQYIPQDIPRL
ADPSTYLGPLHTRERLKIIDFAPLIDKIGARTLSWKGFSSGRTETLYKYLSTIP
TYHLSAIPIPKMAIKKIDIKIRAPLMEGEENYKPGASLVNPTVCRPTMGGLGIL
DLHKESTTLRVMMLRWBNKSPMDQNEDESDJHLFOACTRISIGNHKSLFWE
DNMLQVCPKDVSPNLFKLAMKSRRTNKLNSMLFSPROINIRIHEIYOLGMM
LQHIHLTSSMEDKLEWOTSSGITYTASATYLAQFQSGEPFNDNFMGLCGKANNAGSG
"

gene
complement (72467, .77023)
/gene="OSJNB0062B06.7"
complement {join (72467, .72757, 72908, .73132, 73257, .73392,
74671, .74889, 75097, .75484, 76780, .77023)}
/gene="OSJNB0062B06.7"
/codon_start=1
/protein_id="CAE04049.2"
/db_xref="GI:38345966"
/db_xref="UniProt/TreMBL:Q7XNLI0"
/translation="MRKRADELKEKVRTOIGTGEDIVGTNNLVDAIOHIGIEHLFKOE
IDNTLRIRTSERFSSSLHEBALMFLRLREHGLMSPDVFGKREKRPDDDARLSVTD
HDTRGLSLYNAHLVHGEPELBEALSIARHLSKMSRTDCLPVLANQVRNLNTA
LPRCKRLSTLSIFSEYQOEYSEIILELAKDFNTVQVHLMELKSISBWRDLNTA
YVGLNVARDAVEGYLWSCVLYEKFKDLSFTFPAKMIILVTLMDLDFSDHATIOER
QUNSAIORMDESAVTLPEYLKFRRELRNFKLOQVTDNDKRYATYRKEKOSR
TYVLOEAPRSGDQITLITAMSSVPIELCVSGTGMGVYMTNEFWVASRTYIYASAK
IGRMNDIAMKRGKNGDVASSVECTMNEKHTMEVAIDIKIDSLYDEMRITLQALF
EDHLFPVEEONVLTASMASFYDERDAPTFPLLDITLESLEVNVP1"

Query Match
Best Local Similarity 21.6%; Score 287.8; DB 8; Length 187154;
Matches 439; Conservative 0; Mismatches 87; Indels 49; Gaps 5;

QY 461 GAGCAAGGCTGTTCTGCTGGGCTGGAACAATTGGCAAGGGGACTGGCGAGCATC 520
Db 57943 GTGCGCAGGCTGTTCTGCTGGGCTGGAACAATTGGCAAGGGGACTGGCGAGCATC 57884
QY 521 TCGGGCAACTGTCGATCGCGGACGCAAGGAGGCGGACCGCGCAGAGTAC 580
Db 57883 TCGGGCAACTGTCGATCGCGGACGCGGACGAGGAGGCGGACCGCGCAGAGTAC 57824
QY 581 TTCTCCGCTCACTCCATGAACCGCGACCGCGCGCTCCAGCATCAAGCATCAC 640
Db 57823 TTCTCCGCTCACTCCATGAACCGCGATCGTGGCGCGCTCCAGCATCAAGCATCAC 57764
QY 641 AGCGTCAACCGCGGATCAGGTGCGCGGACGAGGAGGCGCGCGATCACCGCCACG 700
Db 57763 AGCGTCAACCGCGGCG--ACGTTGCGCGGACGAGGAGGCG--CCGATCACCGG 57717
QY 701 GCCACGGGCAACCCCGGCGGCGCGGCGTGGCGCGCGCGGATGAAGACACACAC 760
Db 57716 CCAAGCCCAAGGACGCTGCGCGCGCGCGTGGCGCGCGGATGAAGACAC-- 57665
QY 761 CACCAACCGGCGGCGCGCGCGCGCCATGTCATGAACAGCGCGCGCCATGAGCAC 820
Db 57664 -----CGGCGCGCGCGCGCGCGCGCGCATGTCATGTCGTCATGAGGCGCAC 57618
QY 821 CCGCGTGGCGGACATGTCGCGCGCGCGCGTGGCGACGCGGTCGTCGCGCGCGGCG 880
Db 57617 CCGCGTGGCGGCGGACATGTCGTCGCGCGCGCGCGCGCGCATGTCGTCGCGCGG-- 57560
QY 881 CACCGCGCGTACGTCGTCGCGCGTACCGCGCGCGTCCGCGCAAGATGACACATGA 940
Db 57559 -----CTACCGCGCGCGCGCGCGCGCGCAAGATGACACATGA 57511
QY 941 CGCGCAATGACGAGCATGACAGATTTCTTCTCTCTCTCTTCTGATGTCATCTTGA 1000
Db 57510 GCACACATGATGACATGATGATGATTTCTTCTCTCTCTTCTTGAATG--AATCTTGA 57453
QY 1001 TTGTGTTCTTTGTGATGTCGCGCGCTCATGTCCT 1035

Db 57452 TTGTGTTCTTTGTGATGTCGCGCGCTCATGTCCT 57418

RESULT 6
AK068623

LOCUS
DEFINITION
AK068623 1460 bp mRNA linear PLN 24-JUL-2003
Oryza sativa (japonica cultivar-group) cDNA clone:J013153N23, full
insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AK068623
AK068623.1 GI:32978640
FLI_CDNA; CAP trapper.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
AUTHORS
1 The Rice Full-length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-length cDNA Project Team,
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group; Ootomo, Y., Murakami, K.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Nekawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Nikura, J.,
Ikedo, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusunegi, T., Oka, M., Ryu, R., Ueda, M., Matubara, K., RIKEN,
Kawai, J., Carninci, P., Adachi, U., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishi, Y., Itoh, M.,
Kagawa, I., Kondo, S., Kono, H., Miyazaki, A., Osato, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shitagawa, A., Shiraki, T.,
Yoshino, M., and Hayashizaki, Y.

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
2 (bases 1 to 1460)
Adachi, U., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashidume, W.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirooka, T.,
Hori, F., Hotta, I., Iida, U., Iida, Y., Ikeda, R., Imamura, K.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kawagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
Koya, S., Kurihara, C., Kurosaki, T., Kusunegi, T., Li, C., Lu, M.,
Masuda, H., Matubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
Namiki, T., Narikawa, R., Nikura, J., Nishi, K., Nomura, K.,
Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
Osato, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
Shitagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
Tagawa, A., Takahashi, F., Takaku-Akihira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
Yoshimura, A.

TITLE
JOURNAL
COMMENT
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:ekikuchi@nias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica
rice.

URL: http://cdna01.dna.affrc.go.jp/cDNA/NIAS_Rice_Full-length_cDNA_Project_Team:kikuchi.S.,Satoh.K.,Nagata.T.,Kawagashira.N.,Doi.K.,Kishimoto.N.,Yazaki.J.,

COMMENT

Laboratory of Gene Expression, 2-1-2 Kamonodai, Teukuba, Ibaraki 305-8602, Japan (E-mail: ekikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica rice.

URL: <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Naniki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shieniki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Nagata, T., Kurosaki, T., Kusumegi, T., Lu, M., Mayuda, H., Miura, J., Mizuno, K., Nariawa, R., Nikiura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, K., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishii, K., Nomura, K., Numasaki, R., Ohno, M., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akishi, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

FEATURES

source

1. 1988
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultiVar="Nipponbare"
/db_xref="taxon:39947"
/clone="J01313D01"

ORIGIN

Query Match 14.8%; Score 197; DB 8; Length 1988;

Best Local Similarity 57.4%; Pred. No. 4,3e-12;

Matches 400; Conservative 0; Mismatches 285; Indels 12; Gaps 2;

170 GCGCTCGCGGAGGCTGCGCGGCGGAGGCTGCGGAGGCTCGGAGGACTACGAG 229
Db 585 GAGCGGCTGGGCTGCTGCTGCTGCGGCAAGAGCGTGGCGGAGCTATGACGACTACGAC 644
Qy 230 GCGCTGCTGAGGAGCTGCGCGGCTATGACGCGGCGCGCTCCGCTCCGCGTACGCC 289
Db 645 GACCTCGAGAGAGCTGCTGCTATCGAGCGCGGCTCTGCTCCCTCCCACTACGCGC 704
Qy 290 GGGGAGAGTCTCGCGCGCGCGCGGAGCGGAGCGCGCGCGCGCGCTCCAGAGAC 349
Db 705 G-----CGCGCGGCGCGCGCGCGCGCTCCGCGCTTCAGCTTGACTGGAGCGCGCGC 755
Qy 350 GCGCGAGACCGGCGGAGGAGGCGGAGGCGCGCGCGCGGCTAGAGCGCGGCAAGAC 409
Db 756 GATGACCGCGCGCGCTTGGGTTCAAGGCTCTGCTCACTG--GTGCGGCGGCAAGCGC 812
Qy 410 TGCTTCAAGGCGGAGGAGGAGGCGGAGGCGGATCCCATGAGCGAGAGAGACAGG 469
Db 813 GCGCGCGCGCGCTGACCAAGAGGCGCAAGAGGCGCGCTGCGAGAGAGAGAGACAG 872
Qy 470 CTGTTCTTCTGGGCTGAGCAAGTTGCGCAAGGGGGAAGTGGCGGAGCATTCGCGCAAC 529
Db 873 TTGTTCTTGAAGGCTGAAGAGTACCGGAGAGGTGACTGGAGGAGCATTCGCGCAAC 932
Qy 530 TTTCGTCATCTGGCGGAGCGCAAGCGAGGTGGGAGGAGCGGCGGCAAGTACTCTACCGC 589
Db 933 TTTCGTCAGAGCGGAGCGGAGCGAGAGGTGGGAGGAGCGGCGGCAAGTACTCTACCGG 992
Qy 590 CTCAACTCATGAAACCGGAGCGCGCGCTTCAGCATTCACGATCATCAGGCGTCAAC 649

Db 993 CTCAACTCTCGGCGGAGAGAGAGGCGCGCTTCAGCATTCACGATCATCAGGCGGCTCAAC 1052
Qy 650 GCCGCGATGAGTGTGCGCGGCGGAGAGGCGCGCGCGGATCACCGGCGACAGGCGGAGC 709
Db 1053 CTCCTCGGAGAGACAGACAGGAGCAACCTTCGCGCTGCGCGCGCGCTCTGCTACCGCG 1112
Qy 710 AACCCGCGGCGGCGGCGCTGCGCGCGCGCGGAGTGAAGACCAACACACGAGCGCG 769
Db 1113 CACTCTCTCTCTCTCGCGCGCGCGCGCTCTCGAGAGAGTTGGCGTGTGAGAGGCGAAG 1172
Qy 770 GCGGCGCGCGCGCGCGCGCGCGATGATGACAGCGCGCGCGCGCGCGCGCGCGCGC 829
Db 1173 CTTCCGC 1232
Qy 830 GCGCAGATGTGC 866
Db 1233 GCCCAGGTGAAGATGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 1269

RESULT 9
AK111988
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) cDNA clone:001-030-C10, full insert sequence.
ACCESSION
AK111988
VERSION
AK111988.1 GI:37988651
KEYWORDS
PIL_CDNA; oligo capping.
SOURCE
Oryza sativa (japonica cultivar-group)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

AUTHORS

1 The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Nishii, K., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shieniki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group, Ootomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nariawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Nikiura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Ootomo, Y., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
JOURNAL
Science 301 (5631), 376-379 (2003)
MEDLINE
22752273
PubMed
12869764
2

REFERENCES

AUTHORS

TITLE
JOURNAL
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@niae.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.
PAIS Genome Sequencing & Analysis Group: Ootomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Maeda,H., Mura,J., Munro,K., Narikawa,R., Nishikawa,T., Oka,M., Ryu,R., Sugano,S., Suiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hirooka,T., Hori,F., Iida,J., Imanura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,D., Kishikawa,Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Kiyohisa,Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Nunasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Saeki,D., Sato,K., Shibata,K., Shingawa,A., Shiraki,T., Sogabe,Y., Tagami,T., Tagami-Takeda,Y., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.
Location/Qualifiers

FEATURES
source
1..1795
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultiar="Nipponbare"
/db_xref="taxon:39947"
/clone="J013124L02"

ORIGIN

Query Match 12.9%; Score 171.6; DB 8; Length 1795;
Best Local Similarity 57.4%; Pzed. No. 2.1e-09;
Matches 334; Conservative 0; Mismatches 239; Indels 9; Gaps 1;
68 GCGTGCAGCAGGAGAGAGCAAGAGCGGCTTCGAGAGCGGCTTCGCGGCGCGGCGGCG 127
339 GCGTGCAGGAGAGAGAGCGGCTTCGAGAGCGGCTTCGCGGAGAGAGAGAGAGAGAGAG 398
128 CCGCGCGGAG 187
399 GAGCGGAG 458
188 CCGCGGAG 247
459 ATCCGCGGAG 518

Oy 248 GCGGCATGACGCGGCGCGGCGGCTTCGCGGCTACGCGGAGAGAGAGAGAGAGAGAG 307
Db 519 CAGCAGATCGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 569
Oy 308 CCGCGCGGAG 367
Db 570 GCCCGCGGCGGCTTCACACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 629
Oy 368 GAGCGCAG 427
Db 630 GGTATCCGCTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 689
Oy 428 GAGAGCGCCAG 487
Db 690 GAG 749
Oy 488 GACAAATTCGCGAG 547
Db 750 AAGAAATTCGCGAG 809
Oy 548 CCAAGCGAG 607
Db 810 CCGAGCGAG 869
Oy 608 GACCGCGCGCGCTTCGAG 649
Db 870 GACAAAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 911

RESULT 11
AC145389/c
LOCUS
DEFINITION
Zea mays clone ZMBB0054A01, *** SEQUENCING IN PROGRESS ***, 4
ordered pieces.
ACCESSION
AC145389
VERSION
AC145389.6 GI:48762556
KEYWORDS
HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE
Zea mays
ORGANISM
Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD 1 (bases 1 to 198102)
Britten,B., Nusbbaum,C., Lander,E., Butler,E., Wing,R., Bharti,A.K. and Messing,J. Zea mays, clone ZMBB0054A01
Unpublished
2 (bases 1 to 198102)
Britten,B., Nusbbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Archchbi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Camarata,J., Chang,J., Choepel,Y., Collamore,A., Cook,A., Cooke,P., Corum,B., Develiano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagoos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meidrum,J., Meneus,L., Minova,T., Mengesha,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunhphang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Recte,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schnupack,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Teste,J., Theodore,J., Tophan,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (01-JUN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 198102)
Britten,B., Nusbbaum,C., Lander,E., Butler,E., Wing,R., Bharti,A.K.,

TITLE
JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT

Messing, J., Abouelleil, A., Allen, N., Anderson, M., Anderson, S.,
Arauchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhalef, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., Corum, B., Dearellano, S., Diaz, J.S., Dodge, S.,
Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P.,
Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L.,
Grand-Pierre, N., Hafez, N., Hagopian, D., Hago, B., Hall, J.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P.,
Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J.,
Menais, L., Mhova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunhachang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (20-JUN-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 16, 2004 this sequence version replaced gi:48717652.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/MIT Center for Genome Research

Center code: WTHR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@broad.mit.edu

Bharti, AK and Messing, J: The Plant Genome Initiative at
Rutgers, Waksman Institute, Rutgers, The State University of New
Jersey, 190 Frelinghuysen Road, Piscataway, NJ 08854, USA
(http://pgr.rutgers.edu)

Butler, E and Wing, R: Arizona Genomics Institute, Biological
Sciences West, 448A, P.O. Box 210088, University of Arizona,
Tucson, AZ 85721, USA (http://www.genome.arizona.edu)

Project Information

Center project name: L29613

Center clone name: 54_A.1

Consensus Information

This consensus is derived from a shotgun assembly that has been
manually curated. It is the best representation of the BAC that we
can generate without further laboratory work. The draft assembly
has been edited, and if possible, ends identified by vector as well
as by BAC end sequences, and contigs ordered and oriented. Bases
that are not Ns are either above Q20 or manually edited. This
assembly was performed with phrap. All trace files for this project
are available at the NCBI trace repository
(http://www.ncbi.nlm.nih.gov/Traces/trace.fcgi?). An exact list of
reads used in this assembly are available at
http://www.broad.mit.edu/annotation/plants/maize/randomclones.html.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

* 1 20266: contig of 20266 bp in length
* 20267 20366: gap of 100 bp
* 20367 78634: contig of 58268 bp in length
* 78635 78734: gap of 100 bp
* 78735 140318: contig of 61584 bp in length
* 140319 140418: gap of 100 bp
* 140419 198102: contig of 57684 bp in length.

Location/Qualifiers

FEATURES

source 1. 198102
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZM85C0054A01"
/clone_1lb="CHORI-201 Maize B73 BAC EcoRI Library"

ORIGIN

Query Match 10.6%; Score 140.8; DB 2; Length 198102;
Best Local Similarity 58.5%; Pred. No. 1.1e-06;
Matches 264; Conservative 0; Mismatches 184; Indels 3; Gaps 1;

QY 464 CACAGGTTGTTCTTCTGCTGGCTGAGCAATTTGGCAAGGGGAGCTGGCGAGCATTCG 523
DB 28743 CCAGGCTTTCTTCTTGAAGGAGCATGAAGATAAGGGAGAGGAGCATGAGAACTTCG 28684

QY 524 CGCACTTGTGATCTCTGGGAGCGCAAGCAGAGTGGCGAGCCAGCGGAGAGTACTTC 583
DB 28683 CGCAAGTAGTAGAGAGAGCCGAGAGCGAGAGAGTGGCCAGCCAGCGGAGAGTACTTC 28624

QY 584 ATCCGCTCAACTTCATGAACCGGAGCCGCGCTCCAGCATTCACGACCATCCAGC 643
DB 28623 ATCCGCTCAACTTCGCGGCGAGAGAGAGCGCGCTCCAGCATTCACGACCATCCAGC 28564

QY 644 GTCA--CGCGCGGATCAGGTGCGCGGAGAGGAGCGCGCGCGGATCAGCGGACAG 700
DB 28563 GTCAACTCTCCGAGAGAGAGAGCGCGGAGAGAGCGCGCGGATCAGCGGACAGC 28504

QY 701 GCACGCGGAGAGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 760
DB 28503 AACCTCTCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 28444

QY 761 CACACCGCGGAG 820
DB 28443 CACCGCTCAACTTCGCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 28384

QY 821 CCGGTCGCGGAG 880
DB 28383 AACGACCTCTCTCTGAG 28324

QY 881 CAGCGCGGAG 911
DB 28323 ACCGCTCTCTCTGAG 28293

RESULT 12

AC130602/c 126532 bp DNA linear PLN 02-OCT-2004

LOCUS DEFINITION
Oryza sativa (japonica cultivar-group) chromosome 5 clone B1122D01,
complete sequence.

AC130602
AC130602.5 GI:51556280

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
HMG.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Embryophyta; Streptophyta; Embryophyta; Tracheophyta;
Eukaryota; Viridiplantae; Magnoliophyta; Liliopsida; Poales; Poaceae;
Euphorbiaceae; Oryzae; Oryza.

REFERENCE
1 (bases 1 to 126532)
Chow, T.-Y., Hsing, Y.-I., Chen, C.-S., Chen, H.-H., Liu, S.-M.,
Chao, Y.-T., Chang, S.-J., Chen, H.-C., Chen, S.-K., Chen, T.-R.,
Chen, Y.-L., Cheng, C.-H., Chung, C.-I., Han, S.-Y., Hsiao, S.-H.,
Hsiung, Y.-N., Hsu, C.-H., Huang, J.-J., Kau, P.-I., Lee, M.-C.,
Liu, H.-L., Li, Y.-F., Lin, S.-J., Lin, Y.-C., Wu, S.-W., Yu, C.-Y.,
Yu, S.-W., Wu, H.-P. and Shaw, J.-F.
Oryza sativa BAC B1122D01 genomic sequence

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Submitted (13-AUG-2002) Institute of Botany, Academia Sinica, 128,
Section 2, Academia Road, Nankang, Taipei 11529, Taiwan

[illegible]

gene
45797. .46030
/gene="B1122D01.7"
/note="unknown protein"
45797. .46030
/gene="B1122D01.7"
/note="similar to rice EST AK063607"
/codon_start=1
/product="unknown protein"
/protein_id="AAU90113.1"
/db_xref="GI:53749252"
/translation="MLLVEFOICNFVLLICLDYFVFLVHLIQIRABGRSTINRIGQA
MIGPVSLSKVYKLTQVYINNCIFMAGLME"
complement(56089. .57141)
/gene="B1122D01.8"
/note="unknown protein"
complement(join(56089. .56560,56666. .57141))
/gene="B1122D01.8"
/note="similar to rice EST AK068138"
/codon_start=1
/product="unknown protein"
/protein_id="AAU90113.1"
/db_xref="GI:53749253"
/translation="FMKESYMEVLPAPAHYFVGOAAGGFWLPDRRGGGGAWSOE
NKYFEQALALDRDPERMERVALLPCKTVADMTYDLENDVCTIEAGLVPFRPY
GAAGGGGGGFTLMDGDDPAGLGFRRSCVWGGKRAKGPDOERKGVFWTEBHKL
FLMGLKRYGSGDMWNIISRNFTSRTPQVASHAKYPIRLNSGDKDRRSIHDTTV
NLPPDDHGNPSPPSPVLTAAHSSSSAAVSEQGVAVDGPFPPLRGAGHHEMP
HPVAQVIEAGNSHVAGGRDLDSVLVQMOCGQLMQLG"

Query Match 10.4%; Score 138.6; DB 8; Length 126532;
Best Local Similarity 59.1%; Pred. No. 2.1e-06;
Matches 237; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

Qy 466 CAGGCTGTTCTTGTGGGCTGGACAAGTTCGGCAAGGGGAGCTGGCGAGCATCTCGCG 525
Db 56563 CAGGTTGTTCTTGAATGGGGCTGAAGAAGTACGGCAGAGGTACTCGAGGAACATCTCCG 56504

Qy 526 CAACCTGTCATCTCGCGGAGCGCAACGACAGTGGCGAGCGACCGCAAGAACTTCAT 585
Db 56503 CAACCTGTCATCGAGCGCGGAGCGCGAAGTGGCGACCGCAAGAACTTCAT 56444

Qy 586 CCGGCTCAATCCATGAACCGCGACCGCGGCTCCAGCATCCAGCATCACCAGCT 645
Db 56443 CCGGCTCAATCCCGCGCGCAAGCAAGCGCGCTCCAGCATCCAGCATCACCAGCT 56384

Qy 646 CACCGCGCGAGTCAGTTCGCGCGAGCAGGCGCGCCGATCACCGGCAACGAGCCAC 705
Db 56383 CACCTCCCGAGCAGACACGAGCAACCTCCGCGCTCGCGCGCGCTCCGCTCAG 56324

Qy 706 GGGCAACCCCGCGCGCGCGCTGGCGCGCGCGCGCGCATGAAGACACACACCA 765
Db 56323 CGCGCATCTCTCTCTCCGCGCGCGCTCGGAGAGATTGGCGCTGTCAGCG 56264

Qy 766 CCGCGCGCGCGCGCGCGCGCGCGCATGATGACGCGCGCGCGCATGGGCCACCCCGT 825
Db 56263 CAAGCTTCGCGCGCGCGCGCTGGCGCGCGCGCGCGAGCACCACTTCATGCCCATCC 56204

Qy 826 CGCGGCGCGCATGATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 866
Db 56203 CTACGCCCAAGTGAAGATCGAGCGCGCGCATTCGCACTCG 56163

RESULT 13
AC130612/c 138467 bp DNA linear HTG 13-AUG-2002
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) chromosome 5 clone F0485G09,
*** SEQUENCING IN PROGRESS ***, 4 ordered pieces.
AC130612
AC130612
VERSION
AC130612.1 GI:22212963
KEYWORDS
HTG: HTGS PHASE2
SOURCE
Oryza sativa (japonica cultivar-group)
ORGANISM
Oryza sativa (japonica cultivar-group)

REFERENCE
AUTHORS
1 (bases 1 to 138467)
Chow, T.-Y., Hsing, Y.-I.-C., Chen, C.-S., Chen, H.-H., Liu, S.-M.,
Chao, Y.-T., Chang, S.-J., Chen, H.-C., Chen, S.-K., Chen, T.-R.,
Chen, Y.-L., Cheng, C.-H., Chung, C.-I., Han, S.-Y., Heiao, S.-H.,
Hsiung, D.-N., Hsu, C.-H., Huang, J.-J., Kau, P.-I., Lee, M.-C.,
Leu, H.-L., Li, Y.-F., Lin, S.-J., Lin, Y.-C., Wu, S.-W., Yu, C.-Y.,
Yu, S.-W., Wu, H.-P., and Shaw, J.-F.
Oryza sativa PAC P0485G09 genomic sequence
Unpublished
2 (bases 1 to 138467)
Chow, T.-Y., and Hsing, Y.-I.-C.
Direct Submission
Submitted (13-AUG-2002) Institute of Botany, Academia Sinica, 128,
Section 2, Academia Road, Nankang, Taipei 11529, Taiwan
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 90592: contig of 90592 bp in length
* 90593 90692: gap of unknown length
* 90693 100409: contig of 9717 bp in length
* 100410 100509: gap of unknown length
* 100510 137114: contig of 36605 bp in length
* 137115 137214: gap of unknown length
* 137215 138467: contig of 1253 bp in length.
Location/Qualifiers
source
1. 138467
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="5"
/clone="P0485G09"

ORIGIN

Query Match 10.4%; Score 138.6; DB 2; Length 138467;
Best Local Similarity 59.1%; Pred. No. 2e-06;
Matches 237; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

Qy 466 CAGGCTGTTCTTGTGGGCTGGACAAGTTCGGCAAGGGGAGCTGGCGAGCATCTCGCG 525
Db 26532 CAGGTTGTTCTTGAATGGGGCTGAAGAAGTACGGCAGAGGTACTCGAGGAACATCTCCG 26473

Qy 526 CAACCTGTCATCTCGCGGAGCGCAACGACAGTGGCGAGCGCGCAAGAACTTCAT 585
Db 26472 CAACCTGTCATCGAGCGCGGAGCGCGAAGTGGCGACCGCAAGAACTTCAT 26413

Qy 586 CCGGCTCAATCCATGAACCGCGACCGCGGCTCCAGCATCCAGCATCACCAGCT 645
Db 26412 CCGGCTCAATCCCGCGCGCAAGACCAAGCGCGCTCCAGCATCCAGCATCACCAGCT 26353

Qy 646 CACCGCGCGAGTCAGTTCGCGCGAGCAGGCGCGCGCATGACCGCGCGCGCGCGAC 705
Db 26352 CACCTCCCGAGCAGACACGAGCAACCTCCGCGCTCGCGCGCGCGCGCGCTCAG 26293

Qy 706 GGGCAACCCCGCGCGCGCGCTGGCGCGCGCGCGCGCATGAAGACACACCA 765
Db 26292 CGCGCATCTCTCTCTCCGCGCGCGCGCGCTCGGAGAGATTGGCGCTGTCAGCG 26233

Qy 766 CCGCGCGCGCGCGCGCGCGCGCATGATGACGCGCGCGCGCATGGGCCACCCCGT 825
Db 26232 CAAGCTTCGCGCGCGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 26173

Qy 826 CGCGGCGCGCATGATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 866

Db 26172 CTACGCCAGGTGAAGATGAGCGCGGCAATTCGACGCTCG 26132

RESULT 14

AY519528

LOCUS Arabidopsis thaliana 945 bp mRNA linear PLN 07-FEB-2004
DEFINITION Arabidopsis thaliana MYB transcription factor (At1g49010) mRNA,
complete cds.

ACCESSION AY519528

VERSION AY519528

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana (chale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

REFERENCE 1 (bases 1 to 945)

AUTHORS Qu, L. and Gu, H.

TITLE The MYB transcription factor family in Arabidopsis: A Genome-Wide
Cloning and Expression Pattern Analysis

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 945)

AUTHORS Qu, L. and Gu, H.

TITLE Direct Submission

JOURNAL Submitted (07-JAN-2004) Life Sciences, National Laboratory of
Protein Engineering and Plant Genetic Engineering, Peking
University, Beijing 100871, China

FEATURES Location/Qualifiers

1..945

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

1..945

/locus_tag="At1g49010"

/locus_cag="At1g49010"

/codon_start=1

/product="MYB transcription factor"

/protein_id="AA509998.1"

/db_xref="GI:41618993"

/translation="MESVVAATWSREBEKAFENALALHCVBEETEDQNNKSSVNPVK
ALEBKXHYQILLEDVVAIENGQVPLRYHNRKGLIDEAALATSPANRDSHSGSS
BKKNPSTGSISSNGSGSRAEOERRKGIPTVEEHRIFLGLDFGKDWRSIS
RNPIVSRPTOVASHAOKYFRLNSMNRDRRSIHITTVNNOAPAVTGGGQPOVY
KHRAOPQPOPOQOHHPPTMAGLCMYGGAAPVGQPIIAPPDHMSAVGTVPVLP
MGTNNHHNHLGVAPVAVPVPPLPQHPAPSTNM"

ORIGIN

Query Match 10.2%; Score 135.4; DB 8; Length 945;

Best Local Similarity 70.4%; Pred. No. 1.7e-05;

Matches 181; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

393 ACGACGGCGCAAGAGCTGCTCCAAAGCGGAGAGAGGCGCAAGGCAATCCCATGA 452
Db 350 ATGAGAGAAAGTGTGCTCGAGAGCTGAGCAAGAGAAAGAAAGGATTCATGGA 409
453 CGAGAGAGAGACAGAGCTGTTCTTCTGGGGCTGGAACAAGTGGCAAGGGGGAGCTGGC 512
Db 410 CTGAAGAGAGACATCGGTTGTTCTTTGGTTGGAACAAGTTGGAAAGAGATTGGA 469
QY 513 GAGACATCTCGCGCAACTTCGTCGCGAGCGCAACGACGAGTGGCGAGCCAGCGC 572
Db 470 GAACATTTCAAGAACTTTGATCTCAAGAACTCCAAACAAGTTGCAAGTATGCTC 529
QY 573 AGAAGTACTTATCGGCTCACTTCATGAACCGGACCGCGCGCTCCAGCATTCAG 632
Db 530 AAAAATCTTATCATGAGGCTTAATCGATGAACGAGATGAAGGGGCTTAGCATTCACG 589
QY 633 ACATCAGCAGGCTCACC 649
Db 590 ACATCAGCAGTGTAAAC 606

RESULT 15

AY086906

LOCUS Arabidopsis thaliana 1161 bp mRNA linear PLN 14-APR-2003
DEFINITION Arabidopsis thaliana clone 29302 mRNA, complete sequence.

ACCESSION AY086906

VERSION AY086906.1

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana (chale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

REFERENCE 1 (bases 1 to 1161)

AUTHORS Haas, B.J., Volfovsky, N., Town, C.D., Tzoukhan, M., Alexandrov, N.,
Feldmann, K.A., Flavell, R.B., White, O. and Salzberg, S.L.TITLE Full-length messenger RNA sequences greatly improve genome
annotation

JOURNAL Genome Biol. 3 (6), RESEARCH0029 (2002)

MEDLINE 22088475

PUBMED 12093376

REFERENCE 2 (bases 1 to 1161)

AUTHORS Brover, V., Tzoukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and
Feldmann, K.

TITLE Full-length cDNA from Arabidopsis thaliana

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 1161)

AUTHORS Brover, V., Tzoukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and
Feldmann, K.

TITLE Direct Submission

JOURNAL Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
Malibu, CA 90265, USA

COMMENT

This clone sequence is one of 5,000 Ceres full-length cDNAs made
available to Tigr and Genbank. The following quality assessment of
this set was done by comparison with known proteins: two percent of
the clones are estimated to be 5'-truncated; less than one percent
are 3'-truncated; approximately two percent represent alternative
splice variants, including unspliced introns and spliced exons; one
percent may contain premature stop codons; five percent may have
frame shifts in a coding region. A sequence is considered to be
5'-truncated if it lacks the translation initiation start (ATG). A
sequence is considered to be 3'-truncated if it lacks the
C-terminal end of the encoded protein. Please note that these cDNA
sequences are derived from the W or Lier ecotypes and therefore
may contain polymorphisms when compared to sequences from Col-0.
Genet carried out the library production and sequencing of the
full-length clones. Ceres, Inc. carried out the clustering of the
5' sequences, selection of clones, and sequence assembly.

FEATURES

Location/Qualifiers

1..1161

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/clone="29302"

70..1014

/codon_start=1

/product="unknown"

/protein_id="PAM63950.1"

/db_xref="GI:21555865"

/translation="MESVVAATWSREBEKAFENALALHCVBEETEDQNNKSSVNPVK
ALEBKXHYQILLEDVVAIENGQVPLRYHNRKGLIDEAALATSPANRDSHSGSS
BKKNPSTGSISSNGSGSRAEOERRKGIPTVEEHRIFLGLDFGKDWRSIS
RNPIVSRPTOVASHAOKYFRLNSMNRDRRSIHITTVNNOAPAVTGGGQPOVY
KHRAOPQPOPOQOHHPPTMAGLCMYGGAAPVGQPIIAPPDHMSAVGTVPVLP
MGTNNHHNHLGVAPVAVPVPPLPQHPAPSTNM"

ORIGIN

Query Match 10.2%; Score 135.4; DB 8; Length 1161;

Best Local Similarity 70.4%; Pred. No. 1.6e-05;

Matches 181; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

393 ACGACGGCGCAAGAGCTGCTCCAAAGCGGAGAGAGGCGCAAGGCAATCCCATGA 452
Db 419 ATGAGAGAAAGTGTGCTCGAGAGCTGAGCAAGAGAAAGAAAGGATTCATGGA 478

```

QY 453 CGAGGAAGACACAGGCTGTCTTCTGCTGGGCTGACCAAGTTGGCAAGGGGACTGGC 512
    |||||
Db 479 CTGAAGAAGAGCATCGGTTGTTCTTTGGGTTTGACCAAGTTGGGAAAGAGATTGGA 538
    |||||
QY 513 GGAGCATCTGGCGCACTTGTCTATCTGGCGAGCCCAAGCAGGTGGCGACCGCCG 572
    |||||
Db 539 GAAGCATTTCAAGGAACTTGTGATCTCAAGAACTCCACACCAAGTTGCAAGTCATGCTC 598
    |||||
QY 573 AGAAGTACTTCATCGGCTCACTCAATGAAACCGCGCGGCTCCAGCATCCACG 632
    |||||
Db 599 AAAAATCTTCATCAGGCTTACTCGATGAACCGAGATGAAGGCGGCTTAGCATTCACG 658
    |||||
QY 633 ACATCACCAAGGTCACC 649
    |||||
Db 659 ACATCACCAACTGTGAC 675
    |||||

```

Search completed: July 29, 2005, 21:09:51
 Job time : 5931 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 29, 2005, 21:09:58 ; Search time 4916 Seconds

(without alignments)
3016.130 Million cell updates/sec

Title: US-10-630-636-7
Perfect score: 1641
Sequence: 1 MTSQAATTTTAAAAAAMTR.....GHAPYVVGVPAPPAKMHQ 306

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+g2n.model -DEV=xlh
-Q=/cgn2_1/USPRO.epool/US10630636/runat_27072005_123227_12166/app_query.fasta_1.455
-DB=genbml -QMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOFC=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45
-DOCALLN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10630636@cgn2_1_4200@runat_27072005_123227_12166 -NCP=6 -ICPU=3
-NO MMAP -LARGEOUTRY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Genbml.*
1: gb_da.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pac.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_rgs.*
12: gb_by.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1641	100.0	1330	8	AY151042 Oryza sat
2	1594	97.1	142680	8	AP004611 Oryza sat
3	1494	91.0	1329	8	AK068565 Oryza sat
4	857	52.2	53000	8	AP006233 Oryza sat

5	704.5	42.9	1460	8	AK068623 Oryza sat
6	704.5	42.9	1465	8	AK106166 Oryza sat
7	678.5	41.3	945	8	AY159528 Arabidops
8	678.5	41.3	1161	8	AY086506 Arabidops
9	663.5	40.4	187154	8	OSJN00107 Arabidops
10	648.5	39.5	119091	8	AC016041 Arabidops
11	633.5	38.6	897	8	AY159532 Arabidops
12	633.5	38.6	928	8	AY096571 Arabidops
13	633.5	38.6	1593	8	AY072090 Arabidops
14	497	30.3	1988	8	AK068138 Oryza sat
15	492	30.0	1191	8	AB083027 Glycine m
16	489	29.8	1243	8	AB083028 Glycine m
17	480	29.3	1795	8	AK067964 Oryza sat
18	472.5	28.8	126532	8	AC130602 Oryza sat
19	472.5	28.8	138467	2	AC130612 Oryza sat
20	457.5	27.9	867	8	AY159533 Arabidops
21	457.5	27.9	898	8	BT005473 Arabidops
22	457.5	27.9	1640	8	AK118891 Arabidops
23	456	27.8	892	8	AY088362 Arabidops
24	456	27.8	915	8	AY050976 Arabidops
25	451	27.5	679	8	AY091177 Arabidops
26	449	27.4	897	8	AY159529 Arabidops
27	449	27.4	928	8	BT006122 Arabidops
28	449	27.4	1645	8	AK118135 Arabidops
29	448	27.3	1502	8	AK111988 Arabidops
30	438.5	26.7	1204	8	AF239956 Hevea bra
31	437	26.6	808	8	LES243339 Arabidops
32	428.5	26.1	792	8	AY550308 Arabidops
33	428.5	26.1	823	8	AY091265 Arabidops
34	428.5	26.1	1071	8	AY056180 Arabidops
35	428.5	26.1	1169	8	AY084487 Arabidops
36	428	26.1	1029	8	BT012856 Lycopersi
37	423	25.8	804	8	AY159530 Arabidops
38	423	25.8	63898	2	OSJG00045 Arabidops
39	422	25.7	150206	8	OSJN00083 Arabidops
40	420.5	25.6	834	8	AY151931 Arabidops
41	413	25.2	198102	2	AC145389 Zea mays
42	410.5	25.0	79073	8	AB008271 Arabidops
43	410.5	25.0	83485	8	AB062872 Arabidops
44	403.5	24.6	1745	8	AY077454 Antirrhin
45	399	24.3	134982	8	AP003279 Oryza sat

ALIGNMENTS

RESULT 1
AY151042
LOCUS
DEFINITION Oryza sativa (japonica cultivar-group) transcripti factor MYB51
ACCESSION AY151042.1 GI:24850302
VERSION
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
AUTHORS Lu,C.A., Ho,T.H., Ho,S.L. and Yu,S.M.
TITLE Three Novel MYB Proteins with One DNA Binding Repeat Mediate Sugar
and Hormone Regulation of alpha-Amylase Gene Expression
JOURNAL Plant Cell 14 (8), 1963-1980 (2002)
MEDLINE 22163024
PUBMED 12172034
REFERENCE 2 (bases 1 to 1330)
AUTHORS Lu,C.A., Ho,T.H., Ho,S.L. and Yu,S.M.
TITLE Direct Submision
JOURNAL Submitted (17-SEP-2002) Institute of Molecular Biology, Academia
Sinica, Taipei, Nankang 115, Taiwan
FEATURES
source
1..1330
/organism="Oryza sativa (japonica cultivar-group)"

'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein.
The orientation of the sequence is from M13 to -21M13 of the BAC clone. This sequence of OJ1005 B10 clone has an overlap with P0516004 clone(DDJ: AP003276) at the position 1 to 103,355 of 5' end. The nucleotide sequence of this BAC clone was generated by combining Monsanto and RGP-Japan sequencing data. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://xrp.dna.affrc.go.jp/Genomeseq.html>.

FEATURES
Source

```
1..142680
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="11"
/clone="OJ1005_B10"
join(3399..3512,4006..4227,4265..4328,4613..4792,
6611..6996)
/gene="OJ1005_B10.1"
join(3399..3512,4006..4227,4265..4328,4613..4792,
6611..6996)
/gene="OJ1005_B10.1"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC05637.1"
/db_xref="GI:21902090"
/translation="MGHFLCTGTRAPPSMAVNGREYQPIGRHGLARSAVYRPPRG
GGEGRERRRPLASLGGRERKPEBVMVTPDDCSTDCNLSYFQLPLHLRSSRPA
GQGPWPRPHAEHETCGLKPRVGVAGAGYSFMAYSRGHRLGVGRSLPNCR
EPRLATLPCHLHATIAEDFGRTLDNRKFTYCPSESGAAGARSTASRRRTSSM
VDGTEVGAALFVVRRRRIWPRGPVADLSPRGSGSGLPPEARRWIRPRPABELL
SRPPRAAHSRVPPEPPELAAAFRAHSVEGEGRGREGCFIQP"
join(9727..9823,11568..11674)
/gene="OJ1005_B10.2"
join(9727..9823,11568..11674)
/gene="OJ1005_B10.2"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC05638.1"
/db_xref="GI:21902091"
/translation="MAEAACTRASQOQDEAAGIENGCKQLPEDTTCVAAAGAAPP
SMSNSVALINMGCGSSLPHE"
/gene="OJ1005_B10.3"
complement(13538..14391)
/gene="OJ1005_B10.3"
complement(13538..14391)
/note="gag-pol polyprotein
probably inactive due to frameshift(s) in CDS
probably inactive due to stop codon(s) in CDS
pseudogene"
/pseudo
join(16900..17014,17370..17504,18435..18718)
/gene="OJ1005_B10.4"
join(16900..17014,17370..17504,18435..18718)
/gene="OJ1005_B10.4"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC05639.1"
/db_xref="GI:21902092"
/translation="MVTTSKYGDHSGBELVLGPHPATLLOLLOPILPTGSGIR
KGRKAHNSIMLTASYVTEKKAALOWOHLPTISCHKCGODLGGSGEGCRAGC
PEKAVTTAGDERGERRLAQBVALHAKAEGEKRETRREGEDLTKRRAYQERT
ARTWLRKRTSGSDN"
join(22941..24398,24894..25130)
/gene="OJ1005_B10.5"
join(22941..24398,24894..25130)
/gene="OJ1005_B10.5"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC05640.1"
/db_xref="GI:21902093"
```

```
/translation="MTIFQTLTWMEAAAVACIFISFRLLKODYVDPKRVKODHK
NIRSLNVFYLALAHSMYSYIILVAQPLAVPRVHVSYSKLHGMDSVSTAK
KLMAFIRVREALEMMULIFGKELASDSIDDDKVRILDIYIIRMEKHLSPAGVL
DIYIEYKRVLRVLSIRASTETLERVNMGLNKSLEEBTRHAASIVLELAPYL
FENLPAMPQILASLLTGRERINRPHNDINSVELTCVKTLELLENPNRRSVADA
NVLAKIVELNFRDCKVPVDETERSDSEIVIASLNVLHKLAVSTIGEGELRST
ISKSEISMSTRKLLYQHDNKSLSVAHVILSCLMHEFARVIGSSCQIVKLVS
SVLRPMPIVODGNGSGTVADSAQAVALSTGSEVNRATILDIRINLELVGMLFAS
MEGRIMLHLKLRTYSGPGYQOLKVIDGSPKDLKELEKAEGLDSFGLSSQ
ICHSLQALDPKALQSANLMDTYADQGTERTVOISYHRFPDGKSKNTQSDMAOK"
complement(join(26516..26605,26825..26971))
/gene="OJ1005_B10.6"
complement(join(26516..26605,26825..26971))
/gene="OJ1005_B10.6"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC05641.1"
/db_xref="GI:21902094"
/translation="MENIYTSRVIFKVLGGDRVVISYKIQCLRVLLGDQDDPWYKR
DPQGRILVPYVNTLIIRYSITRQLYTPQCY"
complement(join(28040..28278,28339..28375))
/gene="OJ1005_B10.7"
complement(join(28040..28278,28339..28375))
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC05642.1"
/db_xref="GI:21902095"
/translation="WGAADPAETAAHPABSGVFSRRPQREGRVAAAVRRHG
GSGSVGRSGRRIRRRRATTAATAVAVPRLDPSPTAARTATDLC"
join(30138..30218,31910..32079,32238..32400)
/gene="OJ1005_B10.8"
join(30138..30218,31910..32079,32238..32400)
/gene="OJ1005_B10.8"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC05643.1"
/db_xref="GI:21902096"
/translation="MQLNPKERTRNTKAKGRCARIVLNKNAFTTIDVADGCG
GGGTRDLPRLMRRTETATYAPERSHADCKDGGCGCRADNSGDGLTAVAPAFTE
VVDGSGSYGGRRHMRERKRGREGGEDTDS"
33576..33806
/gene="OJ1005_B10.9"
33576..33806
/gene="OJ1005_B10.9"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC05644.1"
/db_xref="GI:21902097"
/translation="MRRVTQMKSTAAAGADAFAETGDEDDDEDDGVEDDEDD
DGGCGHGBAERGDEDAAGAEVARRGGG"
join(35782..36026,36318..36486)
/gene="OJ1005_B10.10"
join(35782..36026,36318..36486)
/gene="OJ1005_B10.10"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC05645.1"
/db_xref="GI:21902098"
/translation="MAVPPKNDAAFPDPPNDERLPEKXLTNAVRIAFYRVWGC
GALAFNATVVLGGVATLILKODFMFTIIVFEATCCPKFTLTSRLIADTPLDSD
NICVRGVRPSSFTACTDILKLFPMRHHOSRSL"
complement(join(39042..39137,40177..40346,40543..40590,
41257..41417,42100..42248))
/gene="OJ1005_B10.11"
complement(join(39042..39137,40177..40346,40543..40590,
41257..41417,42100..42248))
/gene="OJ1005_B10.11"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC05646.1"
/db_xref="GI:21902099"
/translation="MARTGAAVAHARTDGLATFENPMTGKVKYELGCTSSDEPRAPH
```


/db_xref="taxon:39947"
/chromosome="1"
/clone="B1249E06"

ORIGIN

Alignment Scores:

Pred. NO.: 6.45e-27 Length: 53000
Score: 857.00 Matches: 157
Percent Similarity: 100.00% Conservativity: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 52.22% Indels: 0
DB: 8 Gaps: 0

US-10-630-636-7 (1-306) x AP006233 (1-53000)

Qy 150 ArgLeuPheLeuLeuGlyLeuAspLysPheGlyLysGlyAspTPArgSerIleSerArg 169
Db 63 AGCGTGTCTTCTGCTGGGGCTGACAGTTCGGCAAGGGGAGCTGGCGGAGCATCTCGCG 122
Qy 170 AasnPheValIleSerArgThrProThGlnValIleSerHisAlaGlnLysTyrPheIle 189
Db 123 AACTTCGTCACTCTCGCGACCGCAACGACAGGTGGCGACCGCGCAGAACTCTCATC 182
Qy 190 ArgLeuPheSerMetAsnArgAspArgArgSerSerIleHisAspIleThrSerVal 209
Db 183 CGCTCACTCACTCAAGACCGGACCGCGCTTCAGCATCCAGCATCACCGAGCGTC 242
Qy 210 ThrAlaGlyAspGlnValAlaAlaGlnGlnGlnValAlaProIleThrGlyHisGlnAlaThr 229
Db 243 ACCCGCGCGCATCAAGTTCGCCGCCACAGAGGGCCCGCGATCACCGGCCACAGCCACG 302
Qy 230 GlyAsnProAlaAlaAlaAlaLeuGlyProProGlyMetLysHisHisHisHisHis 249
Db 303 GGCAACCCCGCGCGCGCGCGCTGGCGCGCGCGCATGAAGCACCAACACACACACAC 362
Qy 250 ProGlyGlyAlaProProProMetProMetProMetProMetProMetProMetProVal 269
Db 363 CCGGCGCGCGCGCGCGCGCGCGCATGACAGCGCGCGCGCGCATGAGGCGCGCGCGTC 422
Qy 270 AlaGlyHisMetValProAlaAlaAlaValAlaGlyThrProValValPheProProGlyHisAla 289
Db 423 GCGCGCGCATGTCG 482
Qy 290 ProTyrValValProValGlyTyrProAlaProProAlaLysMetHisGln 306
Db 483 CCGTACGTCGTGCG 533

RESULT 5
AK068623 1460 bp mRNA linear PLN 24-JUL-2003
LOCUS AK068623
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J013153N23, full insert sequence.
ACCESSION AK068623
VERSION AK068623.1 GI:32978640
KEYWORDS FUL_CDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Euphyllophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Eupharatoidae; Oryzaceae; Oryza.

REFERENCE

AUTHORS

1 The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotte, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Ohtsuka, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Nishikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Mura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN;

TITLE

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kawaga, I., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
Science 301 (5631), 376-379 (2003)

TITLE

JOURNAL

Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: kikuchi@nias.affrc.go.jp, tel:81-29-858-7007, fax:81-29-858-7007)
This clone is one of the 28k full-length cDNA clones from japonica rice.
URL: http://cdna01.dna.affrc.go.jp/cDNA/NIAS_Rice_Full-length_cDNA_Project_Team_Kikuchi,S.,Satoh,K.,Nagata,T.,Kawagashira,N.,Doi,K.,Kishimoto,N.,Yazaki,J.,Ishikawa,M.,Yamada,H.,Ooka,H.,Hotte,I.,Kojima,K.,Namiki,T.,Ohneda,E.,Yahagi,W.,Suzuki,K.,Li,C.,Ohtsuka,K.,Shishiki,T.,andYamamoto,M.

COMMENT

FAIS Genome Sequencing & Analysis Group; Ohtsuka, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Mura, J., Mizuno, K., Narikawa, R., Nishikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K., and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashizaki, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Hotte, I., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishikura, K., Nomura, K., Numasaki, R., Ohno, M., Ohtsuka, N., Oka, Y., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sato, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takatsu-Akibira, S., Tanaka, T., Tomaru, A., Toyu, T., Waki, K., Yasunishi, A., and Hayashizaki, Y.
Location/Qualifiers
1. 1460

FEATURES

source

/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultiivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J013153N23"

URL : <http://cdna01.dna.afrc.go.jp/cDNA/>
 NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yasaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Onsenki, K., Shishiki, T. and Yamamoto, M.
 FAIS Genome Sequencing & Analysis Group: Otsomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurotaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariakawa, R., Nishikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Teunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
 Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hangaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa, Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Otsu, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.
 location/Qualifiers
 1. 1465
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultiVar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="001-208-B07"

ORIGIN

Alignment Scores:
 Pred. No.: 8.4e-22 Length: 1465
 Score: 704.50 Matches: 166
 Percent Similarity: 58.57% Conservative: 22
 Best Local Similarity: 51.71% Mismatches: 76
 Query Match: 42.93% Indels: 57
 DB: 8 Gaps: 13

US-10-630-636-7 (1-306) x AK106166 (1-1465)

QY 12 ALaAlaAlaAlaAlaAlaATPrThArgGluaApsApySaIaPheGluaAAlaLeuAla 31
 Db 279 TCCCGTTGCCGCGCGGTGGTGAACCGACAGAGAAGAGCGCTTCGAAACCGCTGGCG 338
 QY 32 ALaCySaAlaAlaProProAlaApsGlyAlaProApsApsApsPheAlaAla 51
 Db 339 ACGGTG-----GGGATGACACAGAGAGAGAGAGCGGCTGTCGAGAG 383
 QY 52 LeuAlaAlaSerValProGlYAlaArgSerAlaGluiGluiValaArgArgHisTyrGlua 71
 Db 384 CTACGAGAGCGCGTGGAGGG--AGACGGCCGACCGAGGTGAGCGCGCCTACGAGGTG 440
 QY 72 LeuValGluaPValAlaAlaAlaIleApsAlaGlYArGValProLeuProArgTyrAlaGlY 91
 Db 441 CTGGTGGAGGACGTGACGCGCATGAGCGCGGCGGCGCTCTGTGTACGCCGGC 500
 QY 92 GluiGluiSerAlaAlaProProApsGlyAla-----GlyAlaAlaAlaAlaAlaSer 108
 Db 501 -----GACGGGGCGCTCGACGAGAGCGCTCTGCGCGAGGTGGAG 539
 QY 109 LysApsGlyGlYHisArgArgApsGluArgLysGlyGlYGlYGlYTYrApsGlyGlY 128
 Db 540 AAGGGCGGTGT-----GGGGAGAGAGGTGGAGGTGGAGGGGG 578
 QY 129 Lys-----SerCySerTylSaIaGluiGluiGluiAArgArgLysGlyTleProTyr 144
 Db 579 CATGGGAGAGAGGGGTGGCTAAGTCTCTGACGAGAGCGCCGGAAGGGATGCTGG 638

QY 145 ThGluiGluiHisArgLeuPheLeuLeuGlyLeuApsTyrPheGlyTyrGlyApsTyr 164
 Db 639 ACGAGAGACAGACAGACGGCGTTCCTCTTGGACTTGAGAAAGTACGCAAGCGCTGG 698
 QY 165 ArgSerTleSerArgApsPheValIleSerArgTyrProThrGluiAlaAlaSerHisAla 184
 Db 699 AGAGATCTCAAGAAACTTGTGATCTCAAGACCCCAAGTAGTACTGATGCA 758
 QY 185 GluLysTyrPheIleArgLeuApsSerMetApsArgApsArgArgSerSerIleHis 204
 Db 759 CAGAGATATTTATTCGCTGAACTCAATGAACAGAGAGCGCGCATCAAGTATACAT 818
 QY 205 AspTlePheSerValThrAlaGlyApsGlnValAlaAlaGlnGlnGlyAlaProIleThr 224
 Db 819 GACATACCGCGCTGAACCAATGAGATACATCTGCTCTAG-----GGCCAAATCACA 872
 QY 225 GluHisGlnAlaThrGlyApsProAlaAlaAlaAlaLeuGlyProProGlyMetLysHis 244
 Db 873 GGT---CAGCCAAATGCC---CATCAGCA-----AATCTGSAANAATCCTCT 914
 QY 245 HisHisHisHisHisProGlyGlyAlaProProProweProweTyrSerAlaAlaPro 264
 Db 915 AAGCAGTCTTACAGCCAGCAAAATGCGCTCCAGCGCTCGATGCTTATGTACGACA--- 971
 QY 265 MetGlyHisProValAlaGlyHisMetValProAlaAlaAlaGlyThrProValAlaPhe 284
 Db 972 ATTGACAGCAGCATGTGTGTCTCTTGTG---TCCGACATGTGACATCTCTGTACACTT 1028
 QY 285 ProProGlyHisAlaProTyrValAla----- 293
 Db 1029 CCGTTCCTGCTGACCTCATATAGCTATGACATGATGATGCCCTGTGCCGTGAGCTGTA 1088
 QY 294 -----ProValGlyTyrProAlaProAlaLysMet 304
 Db 1089 GTCCCTGTGCCCAATGATGATGCTCCCAATGCCCTCAATGCCCATGCCAGCACAACATCT 1148
 QY 305 His 305
 Db 1149 CAT 1151

RESULT 7
 AY519528
 LOCUS
 DEFINITION
 Arabidopsis thaliana MYB transcription factor (At1g49010) mRNA,
 complete cds.
 ACCESSION
 AY519528.1 GI:41618991
 SOURCE
 ORGANISM
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 945)
 QY, L. and Gu, H.
 The MYB Transcription Factor Family in Arabidopsis: A Genome-Wide
 Cloning and Expression Pattern Analysis
 Unpublished
 2 (bases 1 to 945)
 QY, L. and Gu, H.
 Direct Submission
 Submitted (07-JUN-2004) Life Sciences, National Laboratory of
 Protein Engineering and Plant Genetic Engineering, Peking
 University, Beijing 100871, China
 location/Qualifiers
 1. 945
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 1. 945
 /locus_tag="At1g49010"
 1. 945
 /locus_tag="At1g49010"

FEATURES
 source
 gene
 CDS

gene /chromosome="4"
/clone="OSJNB0062B06"
/clone_1ib="CUG1-OSJNB"
3932. .6799
/gene="OSJNB0062B06.1"
/join(3932. .4688,5168. .5282,5395. .5761,5847. .6161,
6241. .6561,6647. .6799)
/gene="OSJNB0062B06.1"
/codon_start=1
/protein_id="CAE04043.2"
/db_xref="GI:38345960"
/translation="MTNLAELYMMVHGVLPLVGIIVEITLHGTCTYPRDFQAIL
PSMNPNSILFTGMOKLEELRKMKRKLAVQGOORFELCDKKGRIYKRKY
OEVLKANGTCHSCAPKRLHRPCTHYIAAEEGIDAVYVSGYPSKQAIYHWS
EIVGIGIAEFTNEDEVTLPDPSKLRGKARRETRIRNDMDSGRVHKSCDC
ERGTYKACPEKDEKPNAAEAGSSAVDGAAPTDGTSYTSVVGEGSRFTVPSYC
GTGSDMDVPGTDVRLREAGLLPMCLVEAAGVDVDPARRMTVRSLLAALVDNRPE
THPHLPGEVAPTLQDLSYLLGLPLAGDAVRPVTVGMQDILTAAPALVORAPHL
LEPLAHNTGTPTKMLQFTVEQLQAEDEYSYRCLAEVLLMFGVMFCGGHGH
YKSYLSMARRRGVEDPNMRSDCPVLLPMPCORPLCLGDRCOLMARSRRDGRH
FRCGNVDERTRTACAYIEWVDTEPNYLDITTCLOEGRWYFASSTEQYQRRAYE
ROCEQSDMRVLTALPPEMARPRCGDRCOVLRSTKPTTLGKRFVVCINTLDDDF
MEPRRCQYRSMIDTRAVLTPESHVQLEPEQYKVTYAREFERGESSRRG"
complement(10043. .18817)
/gene="OSJNB0062B06.2"
complement(join(10043. .10336,10864. .11112,13427. .13562,
16168. .16386,17655. .18033,18577. .18817))
/gene="OSJNB0062B06.2"
/codon_start=1
/protein_id="CAE04044.2"
/db_xref="GI:38345961"
/translation="MEKADKLKQDRIIFETHNTAKGMHVDANQRLGIDHLODEI
NNVSDIKESFTSSLHEVLAHFRLEHGIWSPDVEFKEGSDGFINTIADDER
ALSLYNAAHLVDEPELEBMSAPARHLLEMDRSLKAPLQDQINRAHLPLPT
YKRYEMLHYMLEYGOEEBECIVLLDLAHLLEFNLOHVLKELAFSCQMKDLGYVEL
SHVRDASAYLSYALFYEBNLITRMILAKIIVFIYIMDYDHDITIECRCLNE
AIORWDESAISILPEYMKFYRALQNYFRETEAOEASDKYRVTCKMKEPONLSYVL
OEFEMLHONYKPAKERYALSTLSSTVPLCVTAVGGDAVTKESFELTVRSRAVI
ACAKTRFMNDIAFKSGRKNGKGDANVCEYKMNKRTSVLADKISMSISEKRTL
NOVCRDHQOPVQORLNLNLAVSVPFYDDKKDAYTFYSYIOEIVGDLFVNPVPI"
26861. .27649
/gene="OSJNB0062B06.3"
26861. .27649
/gene="OSJNB0062B06.3"
/codon_start=1
/protein_id="CAE04045.2"
/db_xref="GI:38345962"
/translation="MVPLNDLFLPRADLEYLTOMFSQOELDSULTIPADNVGPPD
ENGFFATKCHITIAHFABEATDLOVLANSYTLVFKSPSTVDFRISLMSGLK
LTKMLADRLQVILLEVNGQYEFIKGTIOECLAQAEYIHQCOQSREIITILKDR
EKADDTIHTAITSYVENNGFPOKMLNVAQFASASTIILNGVSGSFPCKGKGV
GDPSPILFVLGAELIQRIVRAFNILGKLPTVYGLPRVLDDQK"
complement(36241. .36855)
/gene="OSJNB0062B06.4"
complement(36241. .36855)
/gene="OSJNB0062B06.4"
/codon_start=1
/protein_id="CAE04046.2"
/db_xref="GI:38345963"
/translation="WPTTYAARDOBELARRNGRPEESSTRNGKDDGDTGCVGEAR
PVTPLRCRRRRRRRPAHGHGRCGRGCRDRKGGTGGGLRTNLRAGMGGRRR
GAGGDDGVRRSGAATGRLEBAAGMGRCGRGCRCHVLRKLRKRENGRRGPHFI
GAGREPATGRAGKGGNGRSGWELGIEIEEGKIEFPLINLE"
complement(45605. .50168)
/gene="OSJNB0062B06.5"
complement(join(45605. .45901,46925. .47173,47477. .47612,
48065. .48283,49052. .49430,49925. .50168))
/gene="OSJNB0062B06.5"
/codon_start=1

gene /protein_id="CAE04047.2"
/db_xref="GI:38345964"
/db_xref="UniProt/TREMBL:Q7X6A7"
/translation="MVRAEKLKQDVRTMPTCTCSTGRLOLVNADIGLIDHLPKEE
IEYLSINMSEPTSSLDHVALRFRLLRGRHVS PDVKNFKPGDORFVSGTINPD
RGLSLYNAAHLTHDEPELEAISFTQHLASISGDLNPLHIDQINRLDVLPLR
TYRMEETLCYMPREYROEGHIPLILEAMDENLQYHVAHLEKLAISMMWDLGYMG
LSYIRDRVSEYVWSYVVEYEDSALARMIPETKIAPIILMDDYDSYATIOEGRKLN
BAIORMBESATAPLPEYIKKPYSAIILTKPFEELIYVGGOGYRIDHTKAFONLSAY
LOEAMSYQNTKPSPEEOVALSTYSTVPLCSYTTVGGRCALTNFAEAMNIDGAK
IACAKTRFMNDIAFKRGRKNGKGDVSTVECTYKNEKVTSEBGLFKIDIMDEMR
TINPLCEHRLFAVOQVNLALCAFFGKRDATYFSTHLOETVESLFVRVSI"
complement(64764. .71118)
/gene="OSJNB0062B06.6"
complement(join(64764. .64805,64958. .65940,66221. .67084,
67774. .67858,68435. .68536,68753. .69149,69306. .69546,
69909. .70953,71089. .71118))
/gene="OSJNB0062B06.6"
/codon_start=1
/protein_id="CAE04048.2"
/db_xref="GI:38345965"
/translation="WTVLRYNFPAPAPADVFNSSAVTFKQGLSKSTAGVSVH
KRLQPVARDSLPEPGPSKVSASVYLSKSRVIVSHPPEDASVLLPSMSAMSEVKAK
YMKRKHSSRSSTLEERTRLRAHPLHVMGRCFCLSGSHRVLDLYLLSVGSEV
LDMKSSAPRDLALNPOMLLPYPLCFQIOISPHSLVQASLLROASRRRPMMS
WCAGOTPILILPLGRSSATMRDLICVPLMRGMLTRREGMRAKPLIMTRTAURLH
RRDVQVWKYTPEDFSVTLSPRADAHILCOPRDLCHSGRSTHLKARDROGAHYTR
YRARLCEGILPMHARTEVAVALIGRPVLCGGKKKILMEVVAHILAIIAAPSPGV
VSRAGATLTVGSMIKASTTRTTVAATKPLPEPAGCALVPELEFKFLVPLQVR
VOHSSSPSHRLKPTQAAVINS PCNVKRPKHSKPRGAGCALVPSPSMELFLHQ
DNGCONHOLDPMLBAVINYS PDGPEOQSSPYVAPTEGLAVQSVNVAASDLIA
SLATPRKQTPRSLIADIVKGGCKKILVAGRRNFPPLPDGTREGGILAA
LDNYRRLSQSETRNLSPEBEALHRSIKAKITGLALNRYITRORSULTYKEDGAT
KLPHIKANARKKQFIHYLOSTFPAISHQKEBULFANRYSIKRQSRSLWSQOS
LGMSPLSDOEKALIMEBELITKSLSPSKAPGPDFTSAFYKCDLIGHDLFNAY
RSLFKLQNOQDIDNATCHI CILPKKQASAPENHRPILSIPAKIISKLANLAPR
MAELVTNOSAFIKRAIHDFELVQNVLYKFRKKPQTLMMKLDISRAPVANVMA
MEVSLVADQEISIKRQNLNPGASGLQNNVHTOITLPIACGLANVOYIPDILPRL
ADPTSYLGLPIHRELRKIDPAPLIDIKGARTLSMKGKFFSGRETIVYVYSIP
TYHLSAIPFIPMAIKKIDKIRARFLMGEENPEPASLVNMPTECPKTMKSGIL
DLHREFTTLRVRWMLMWDNSKPMQMBEPCDSIDLHPACTRISIGNHKSLFWE
DNMLQGVCPKQVSNPLFKLARNKSTYKELLSMYSFSGFOITNIDQIHBLVQGM
LOHILHLSNMDKILEWQTSGIYTAASAVLAQFSGSEFNDNFMGGLGCVKAMNAGSG
"
complement(72467. .77023)
/gene="OSJNB0062B06.7"
complement(join(72467. .72757,72908. .73132,73257. .73392,
74671. .74889,75097. .75484,76780. .77023))
/gene="OSJNB0062B06.7"
/codon_start=1
/protein_id="CAE04049.2"
/db_xref="GI:38345966"
/translation="MRRADLEKERVTOIGTCDI VGTMLVDAI OHLGIEHLFKOE
IDNTLRDIRTSEPTSSLSLHVALMFRLLREKGLTGVDSVDEKGFEPDGDARLSEVID
HDTRGLSLYNAAHLVHGRELEBEAISARHRKASTMRCODLNPVLANQNRALNTA
LPTCKRLSTSLFTSYEGBGYSKILLEBLAKDFNIVONVHLELSISMMWDLGYMG
VYGLNVARDRAVEGYLWSCVLYEKDLSFTPTFAKMLLVATLMDDTFDSATLIOEGR
QUNSAIORWDESAVTLPEYLLKFFRNLNFKVLAQOVDNDKYRATYTRKEFOKLS
TYVIOEABPSFGQDITLAMSVPILCLSGTGVMGVYTMETTFEWSRSTTAIVASAK
IGRFMNDIAFKRGRKNGKGDVASSVECTYKNEKVTSEBGLFKIDIMDEMR
EDHKLTFVVEQVNVLTASMASFYDERKDATTFPLLDOTIESLFVNPVPI"

Alignment Scores:
Pred. No.: 2,84e-18 Length: 187154
Score: 663.50 Matches: 153
Percent Similarity: 65.00% Conservative: 3
Best Local Similarity: 63.75% Mismatches: 13
Query Match: 40.43% Indels: 73
DB: 8 Gaps: 8

US-10-630-636-7 (1-306) x OSJNB00107 (1-187154)

[illegible]

/db_xref="GI:7770353"
 /translation="WSANEPLSSAQAFQEOFLGCFVSRKLLLNHPDHNTQRAFAVA
 SPLITHENNLGNVWMLLSILICGIIICCLGHTYIIRCLARSTFRFMSBPVSLSTR
 GSSNGIKKKALRMPVSVSREPMILPGLDEBVCISLDFVSGQLRLVPLPCNHFVY
 RIDIKMLQOOLCPKRCNCLVETCOCKLIGDSQADSVAAETELVITVITVLEPTELV
 IYVILALPSTELVITVIMIALPREGCEVDTNMYRKYAFNKLIMDEBLCRPSYSDNIS
 YTEWVILCALVCSLGANSTIRCALRCSLVNYSBAGQDVPRLNTGYKRAKLSFOT
 VSYSTWMLPGLDEBCALCLSEFAVEEYVNLPTCHNGHFHRCIDKCLIQTCERKID
 CSQTSLSNLTQPPQDSIIILQIALPPEHWMIFR"
 complement|join(16947..17596,21011..21320)|
 /note="putative RING finger protein gb|AAD08934.1; similar
 to BGT gb|N37302.1"
 /codon_start=1
 /evidence=not_experimental
 /product="P27J15.3"
 /protein_id="AAF69722.1"
 /db_xref="GI:7770352"
 /translation="MAANEILASSVQAQEOBQSLGCFVSRKLLLNHPDHNTQRAFAVA
 BSPILTHENNLGNVWMLLSVILICGIIICCLGHTYIIRCAFPRFMSBPVSLSTR
 CAQSLQEOFLGCFVSRKLLLNHPDHNLPAFAVAPSLITHENNLKGNVWMLLSVIL
 CGIIICCLGHTYIIRCAFPRSSRFMSBPVSLSTRSSNGIKKKALRMPVSVSR
 ENALGIGEEVYICLDFVSGEQLRLKCNHGFVRCIDKMLQOHLILCPKRCCLVE
 TCOKLIGDSQADSVASPTESVIRIDPLREGVNTFRSS"
 complement|join(26020..26439,26567..26652,26749..27750,
 27852..28004,28099..28276)|
 /note="putative two-component response regulator
 gb|AAD12696.1"
 /codon_start=1
 /evidence=not_experimental
 /product="P27J15.4"
 /protein_id="AAF69721.1"
 /db_xref="GI:7770351"
 /translation="MLVKGISGEYEDNTLEBETSEITLSQFPQNTVNLVNTNFT
 TLNKLQIMKQAYAVQSIETDQEKALFLTSCKEHINVIDPHNPGIDGLQALKSIT
 SKLDPVYIMSDNQTESYMKATFYGACDYYVYKPEKESVMAINTQHYRKLITKRPV
 APVSDSPKRSRLDQVADFKIVDEPITNPLITVETEIQPOQSLVQANKFDVQ
 NGYSNIDPNSRKLQVADFKIVDEPITNPLITVETEIQPOQSLVQANKFDVQ
 KATAPOLMTVEISIQVOSGLVQANEFKNGVQSQNMQDSFMSATNPRLTNELD
 QVQSDVLQVSNESQFSDYSQIMNEDNFMNAKKPRMTVEFQPVOSHLVPTDGLD
 RHPSITNTNGNGSIQNMKEKQGRPKRPMTWTEELHOKFLALETIGIGEKANPVY
 LVYECQENRIBGITRSNVASHQKRRINLBERQITPOQIQGNGKATAYOTTLAPSLQSD
 NVNTTIPSTLNGPATTNLQIQNOYQNGFLTNANNQIITNPPPLPYDHHHQOQDS
 SQQFYNLWNNELDLQASGISATDELTYPSLPYDQEYLINGYNN"
 join(32180..33399,32476..32593,32675..32733,32808..32872,
 32944..33037,33113..33193,33279..33390,33706..33793,
 33963..34029,34301..34500,34578..34667,34798..34866)
 /note="similar to MAP/ERK kinase gb|4505153"
 /codon_start=1
 /evidence=not_experimental
 /product="P27J15.5"
 /protein_id="AAF69696.1"
 /db_xref="GI:7770326"
 /translation="WLDYTIASKLSSEITSTVTLAKKLTGEEAVNAKCPULSLANR
 LADCLNNELEFSSVDHPVIRLLHVSQDDDFLVNVLVYCGGTLSTYIQRYGVEED
 IAKFMKQYGALEIYIDNHAIIRDLKBEMLIDSGDGLVLTADSLYAKLRHGX
 LKTVQSPYGALEIYIDNHAIIRDLKBEMLIDSGDGLVLTADSLYAKLRHGX
 SVALPFSRLILQOMHPDICDVCSDALINPATGTGIDEPFLGRIKNSRWVQDTTS
 SQOHTREKAVATINGRLOPFLQDSCDALYQRLTENLWYINTYTEMARVQPTSYLIS
 RSESTIKVLSQLRKAQKQTFPQMLRSWNLNLEMLAKVLEATGTFTPIATLDPASLIS
 "
 join(35184..35343,35430..35545)
 /note="similar to serine/threonine kinase 9 gb|AAD28798.1"
 /codon_start=1
 /evidence=not_experimental
 /product="P27J15.6"
 /protein_id="AAF69697.1"
 /db_xref="GI:7770327"
 /translation="MAPTKGGKTSAAESTVTSSSFPCTLRLTLTPSSVAITTHAK
 PQSKASITDVSDEAVQVIDAPARDGENALLETMSVYFLFR"
 join(37751..37837,38154..38374,38475..38655,38792..38949,
 39065..39485,39625..40145)
 /note="protein kinase-like protein emb|CAB41172.1"
 /codon_start=1

```

/evidence=not_experimental
/product="F2J15.7"
/protein_id="AA669698.1"
/db_xref="GI:7770328"
/translation="MEGSEDASAIIVEPDPPEVLIEDPTCRATIRGEDEVGIVANNOV
RIDDLASPDCELRLEISFVRLIKSLKHIIIFPNYNIIDDKRTVNIITELPTSGIR
OYRKIRKRNKMAVKKCMAROIILTGKLYLSODPPIIHRIOKCDNIFINNHGVYICD
LGATVMEQANASVIGTEPFMAPELVDENYELAIYSFGNCKEMVTFEYPCBR
NSAOIKYKVSQSPKASLRSKVCDPEYMKPIEKCLLPASRLSABEILLDSFLANGLV
MNPDLPLDIPVMPKESFGECRLMSGEPNANRNTKMLSDNNLPVIVSSNGTV
CIVERRAKGNFPLVKGENSENSVSLIRIVDENGRVNIHPLFPOEGDTASVASE
MGEOLTELQNKYEIETLIDVLLYNIIPMKTDVANDHIIHQONQNSKDNHONQASS
OAGSISHSLSSDYCPRSDDEANPTVAATTEDOEAKPSPSLSEEBDEDLKSELKIE
ERREREMKEITRRREBAETLTKRFFEKMOOVE"
/complement(join(40641..40798,40897..41041))
/note="hypothetical protein"
/codon_start=1
/evidence=not_experimental
/product="F2J15.8"
/protein_id="AA669720.1"
/db_xref="GI:7770350"
/translation="MDNKAKTAKAASPSVTKDTPDNLKDKOMTTEKISDAMEM
LEPESKSVDSGSSDESPYSFGTPEVFGGGGGLFVGGRFGGRCGGGFRRC"
join(41349..41637,41888..41922)
/note="similar to NADH-ubiquinone oxidoreductase 12 kd
subunit precursor sp|Q030015"
/codon_start=1
/evidence=not_experimental
/product="F2J15.9"
/protein_id="AA669699.1"
/db_xref="GI:7770328"
/translation="MGRKKGLPEPESAPDGDPEPNYKDPVAMVEMREHIVREKMIH

```

Db	86508	GGGATTCAGTACTGAAAGACATCGGTTCTTCTCTCTTCAATCATCAT	86449
QY	149	-----	149
Db	86448	CTATCCATTGATATATTAGACTTAAGAAGTGAAGAAAGATTGGAGATTCTTACC	86389
QY	150	-----	
Db	86388	TTTATTGAATCTCTGTTTGGTGGAGGTTGTTCTTTGGGTTGGACAAGTTTGAG	86329
QY	161	LYSGIYASPTTPARGSERILESERARGASPHVAIIIESERGRHPRORHNGIYVLI	180
Db	86328	AAAGAGATTGAGAAAGCATTTCAAGAACTTTGTGATCTCAAGAACTCCACACAAAGTT	86269
QY	181	ALASERHISALAGINLYSRPHEILEARGLEASERMETASARGASPARGARGATG	200
Db	86268	GCAAGTATGCTGCAAAAATTAATCTTCATAGGCTTAATCCATGATGACGAGATGAGACGG	86209
QY	201	SESERILEHISAPPIETHRSERVALTHRALGIYASPIYVLIALAAGLNGINGIY	220
Db	86208	TCTTGACATTCACGACATCACC-----ACTGTGACATCAAGCTCTCGGTTACAGGA	86155
QY	221	ALAPROLEIHRNGIYHISGINALATHRGIVASPHRALALALALEUGIYPROPRO	240
Db	86154	GGAGACACAAACCGCAAGTGTTAAACATGACACGACTGAGCCACACACACACACAA	86095
QY	241	GIYMETLYSHISHISHISHISHIEPRONGIYGLIYALAPROPROMETPRO-----	258
Db	86094	CCGACACCAACAAACATCAT-----CCCCAACATGCTGTGATTA	86053
QY	259	---METYSERIALAPROMETGLIYHISPROVALALAGIY-----HISMETVAL	274
Db	86052	GGGATGATAGTGTCGGCCAGTGGGACCAACCATCATGACACACCTGATATATG---	85996
QY	275	PROALALAVAGIYTHPRPROVALPHPRPRO-----	286
Db	85995	GCTTCAGCTGTGGAAACCGTGATCTTCCACCTCCAAATGGAACTCATCATCATCAC	85936
QY	287	-----GLYHISALAPROGYVALVALPROVALGIYRPROALAPPRO	301
Db	85935	CATACCATCATCTTGGAGTTGCTCTTATGCTGTACCG---GCTTATCCGATACCGCA	85879
RESULT 11			
AY519532	897 bp	mRNA	linear
LOCUS	AY519532	Arabidopsis thaliana MYB transcription factor (At5g08520)	mRNA,
DEFINITION		complete cds.	
ACCESSION	AY519532		
VERSION	AY519532.1	GI:41619007	
KEYWORDS			
SOURCE			
ORGANISM		Arabidopsis thaliana (thale cress)	
		Arabidopsis thaliana	
		Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
		rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
REFERENCE			
AUTHORS		Qu, L. and Gu, H.	
TITLE		The MYB Transcription Factor Family in Arabidopsis: A Genome-wide	
		Cloning and Expression Pattern Analysis	
JOURNAL		Unpublished	
REFERENCE		2 (bases 1 to 897)	
AUTHORS		Qu, L. and Gu, H.	
TITLE		Direct Submission	
JOURNAL		Submitted (07-JUN-2004) Life Sciences, National Laboratory of	
		Protein Engineering and Plant Genetic Engineering, Peking	
		University, Beijing 100871, China	
		Location/Qualifiers	
FEATURES			
source		1..897	
		/organism="Arabidopsis thaliana"	
		/mol_type="mRNA"	
		/db_xref="taxon:3702"	
gene		1..897	
		/locus_tag="At5g08520"	

[illegible]

Db		661	GGAGGAACAATCAGCCACG-----CAAGCGTCTCTCAAGCACACACTGTGA	708
Oy		257	MefProMetYrSeSaIaAlaPROMetGIYhiSProValaIaGlYhiMeetValProIaA	276
Db		709	CCTCTATGTATNGAMACACGCCGCATAGTCAGCCA-----	744
Oy		277	AlaValaGIYThrProValaIaPhnePro---ProGIYhiSaIaProTYrValaVal-----	293
Db		745	GCAGTTGGAAACCAGTAACCTCCACACTCCACCTCACTGACATGAGGATTAGATCATCG	804
Oy		294	---ProVala----GLTYrProalaproProalalySMethis	305
Db		805	GCTCCAGTTCCTCGCTCAGTGCTTCTGTGTCAGCAATGAC	846
RESULT 12				
LOCUS	AY096571	928 bp	mRNA	linear
DEFINITION	Arbidopsis thaliana unknown protein (At5g08520) mRNA, complete cds.			
ACCESSION	AY096571			
VERSION	AY096571.1			
KEYWORDS	FLI CDNA.			
SOURCE	Arbidopsis thaliana (chale cress)			
ORGANISM	Arbidopsis thaliana			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosoid II; Brassicales; Brassicaceae; Arabidopsits.			
AUTHORS	1 (bases 1 to 928) Yamada, K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldenlith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A. Arbidopsis Open Reading Frame (ORF) Clones Unpublished			
TITLE	2 (bases 1 to 928)			
JOURNAL	Yamada, K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldenlith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A. Direct Submission Submitted (16-Apr-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA			
REFERENCES	The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN Arabidopsis Full-Length CDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai, J., Hayashizaki,Y. and Shinozaki,K.			
COMMENT	The Saik, Stanford, PGEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldenlith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.			
FEATURES	Annotation is based on the January 2002 version of the Arbidopsis genome submitted to Genbank. Location/Qualifiers			

```

source
    1..928
        /organism="Arabidopsis thaliana"
        /mol_type="mRNA"
        /db_xref="taxon:3702"
        /chromosome="5"
        /clone="UI9230"
        /ecotype="Columbia"
        /note="This clone is in pUNI 51."
gene
    1..928
        /gene="At5g08520"
cDS
    1..897
        /gene="At5g08520"
        /evidence=experimental
        /product="unknown protein"
        /protein_id="PAM20221.1"
        /db_xref="GI:20465999"
        /translation="MTVEEVDGVSVMREDIDAFERALANNTESEBREMKEIADVPC
            KSVQIKEHYELIVEDYTRIESGCVPRLPA GSPSGNAGDEGASAKGSNGSHAGES
            NQAKSKSDGRKRGKIATMEDERHLFLGLDKTYKGKMRISRRFPVTPRPYQASHNA
            OKFTRLSNMVKDRRSIHDIYSVGADVSTPGPLTTGGVNNSNNNNNNNSPPAA
            AGCGFNKAQVAQAPCPMPYGTPALIGQPAVGTVNLPA PPHMA YGVHAPVGSVV
            PGAA MNIGOMPMPRPRTAHR"
        /cds_start=898..928
        /gene="At5g08520"

3' UTR

ORIGIN

Alignment Scores:
Pred. No.:      5,69e-19          Length:      928
Score:          633.50           Matches:     155
Percent Similarity: 58.60%       Conservative: 29
Best Local Similarity: 49.36%    Mismatches:   81
Query Match:     38.60%         Indels:       49
DB:              8              Gaps:          14

US-10-630-636-7 (1-306) x AY096571 (1-928)

Oy      8 ThrThrThrThrAlaAlaAlaAlaAlaTTrpThrArgGluAspArgLysAlaPheGlu 27
      |||
      :|||
Db      4 ACGGTGAGCAAGATTAGTAGTGTCCTGCTGGAGTGGAGATGATGATTCCTTGAG 63
Oy      28 AsnAlaLeuAlaAlaCySAlaAlaProProProAlaAspGlyGlyAlaProAspAsp 47
      |||
      :|||
Db      64 AGACCTCTA-----GCCATAATACCAGTAATCAAGAGAA 99
Oy      48 TrpPheAlaAlaLeuAlaAlaSerValProGlyAlaArgSerAlaGluGluValArgArg 67
      :|||
      :|||
Db      100 CGGTGGGAGAAGATTGCTGCGACAGCTCCAGGC---AAAAGTGTGAACAATATAAGAA 156
Oy      68 HisTrpGluAlaLeuValGluAspValAlaAlaIleAspAlaGlyArgValProLeuPro 87
      |||
      :|||
Db      157 CATTACGAGCTTTAGTTGAAGATGTTACTAGATTGAATCATCAGGATGTCCTCTTCT 216
Oy      88 ArgTrpAlaGlyGluGluSerAlaAlaProProAspGly--AlaGlyAlaAlaAla 106
      :|||
      :|||
Db      217 GCCATAGGCTCTCCTGAAGAGATC-----AATGCCATGTCTGTGATGAAGAGACA 267
Oy      107 AlaSerLysAspArgLysGlyHisArgArgAspGluArgLysGlyGlyGlyTyTrp 126
      :|||
      :|||
Db      268 AGTAGTAAGAAAGAGGTAAACATC-----GGCGAGAGCTTAACCA 312
Oy      127 GlyGlyLysSerCySserLysAlaGluGlnGluArgArgLysGlyIleProTrpThrGlu 146
      |||
      :|||
Db      313 GCAGGTAAA-----TCAAAGTCCGATCCAAGAACGAAAGGATATCGCTGACAGAA 366
Oy      147 GluGluHisArgLysPheLeuLeuGlyLeuAspLysPheGlyLysGlyAspTrpArgSer 166
      :|||
      :|||
Db      367 GATAGCAACAGGTATTCTTCTTGTTGATTAAGTACCGGAAAGGTGATGGCGTAGC 426
Oy      167 IleSerArgAsnPheValIleSerArgTrpProThrGlnValAlaSerHisAlaGlnLys 186
      |||
      :|||
Db      427 ATTTCGCACTTTGTGATGACAGAACCCGACCAAGTTGGGAGCCATGCTCAAAG 486

```


Db		760	GCCATAGGTCCTCCGAAGGATCG-----AATGGCATTGCTGGTGAACGAGA	810
Oy		107	AlaserlysaarglYglYhiisargarabgltuarlglysglyglYglYglYtyrTrp	126
Db		811	AGTAGTAAGAAAGGAGGTAAAGCTCAT-----GCCGGAGAGCTTAACCA	855
Oy		127	GLYglYlysSerCySSerLyalaGlugIngiwArgArGLYVGIilaProtrPhrglu	146
Db		856	GCAAGTTAAA-----TCAAAGTCGATCAAGAACACAAAGGGATATGCCTGGACAGA	909
Oy		147	GLUGluhiisArgLeuPheLeuLeuGLYleuSaplyspheglYtYglYASPrTPraSser	166
Db		910	GATGAGCACAGATTATTTCTTGCTTGTTGATTAAGTAAGGAAAGTGATGGCGTGAC	969
Oy		167	IleserArgaenPheValIileserArgrThrProthrGlnValaIleserHislaGlulys	186
Db		970	ATTTCCTCCAATTGTGTAGTAACAAGAACACCGCAAGTTGCAGCCATCTCAAAG	1022
Oy		187	TyrPheIIeaRgleuAsnSerMetAsnArgAspaARGaAgSerSerIIehIsaplle	206
Db		1030	TATTTCATTCCTGTAAATTCATAAGAACAAAGCAGAACGCCATCAAGATTCACACACTC	1089
Oy		207	ThisSerValThrAlaagLYASPQlnValaIlaIngInglYAlaProclIethrGlyhis	226
Db		1090	ACTAGTGTGGCCACGCGAGAT---GTCTCACGCCACAAAGA---CCATCACTGTGTCG	1144
Oy		227	GlnalathrglyAsn-----ProAlaIlaIlaIla	236
Db		1144	AACAACACCAATYAACAACAACAACAACAACAACAGTTCTCTGCTGTCTGTA	1203
Oy		237	LeuGlYProProGluMetLylshihishihishihishierProGluYglYAlaProtrPro	256
Db		1204	GGAGGAACAATACGCCAAG-----CAAGCCGTCTTCAAGCACACACTTGA	1251
Oy		257	MeCPrometCYrSerAlaAlaPrometGlyhisProvalaIlaGlyhisMetValProla	276
Db		1252	CCTCTTAATGATGAACAACCCGCCATAGSTACGCCA-----	1287
Oy		277	AlaValaGlyThrProValaIalphePro--ProGluhiisAlaProTYrValVal	293
Db		1288	GCAGTGTGAACCCAGTAGAACCTCCAGCTCACCTCAATGCGTTTANGAGATTGATGCG	1347
Oy		294	---ProVal---GlyTYrProAlaProProAlalySMethHis	305
Db		1348	GCTCCAGTCCCTGGCTCAGTGGTTCCTGGTGCAGCAATGAAC	1389
RESULT 14				
AK068138			1968 bp mRNA linear PLN 24-JUL-2003	
LOCUS			Oryza sativa (japonica cultivar-group) cDNA clone:J03135D01, full	
DEFINITION			insert sequence.	
ACCESSION			AK068138	
VERSION			AK068138.1 GI:32978156	
KEYWORDS			PL1_CDNA, CAP trapper.	
SOURCE			Oryza sativa (japonica cultivar-group)	
ORGANISM			Oryza sativa (japonica cultivar-group) Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliopsida; Liliopsidae; Poales; Poaceae; Ehhartridoalee; Oryzaceae; Oryza.	
REFERENCE			1	
AUTHORS			The Rice Full-Length cDNA Consortium, National Institute of Agrbiological Sciences Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashiri,N., Doi,K., Kishimoto,N., Yarakai,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Nemiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Ii,C., Ohtsuki,K., Shienhiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group.; Okomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzaki,Y., Tsunoda,Y., Kuratsaki,T., Kodama,T., Maesuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niihara,J., Ikeda,R., Ishibiki,J., Kawamura,M., Yoshimura,A., Miura,J., Kuenmegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN, Kawai,J., Caminici,P., Adachi,J., Atawaa,K., Arakawa,T., Fukuda,S.,	

TITLE	JOURNAL	REFERENCE	AUTHORS
Yoshino, M. and Hayashizaki, Y.	Science 301 (5631), 376-379 (2003)	2 (bases 1 to 1988)	Adachi, J., Aizawa, S., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Hotate, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, S., Kojima, K., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsumura, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Nariakawa, R., Nishikawa, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Oosato, N., Oga, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.
Direct Submission	Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: ekikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)		
COMMENT	This clone is one of the 28K full-length cDNA clones from japonica rice.		
TITLE	URL : http://cdna01.dna.affrc.go.jp/cDNA/NIAS_Rice_Full-length_cDNA_Project_Team_Kikuchi_S.,SatoH_K.,Nagata_T.,Kawagashira_N.,Doi_K.,Kishimoto_N.,Yazaki_J.,Ishikawa_M.,Yamada_H.,Ooka_H.,Hotate_I.,Kojima_K.,Namiki_T.,Ohneda_E.,Yahagi_W.,Suzuki_K.,Li_C.,Ohtsuki_K.,Shishiki_T.andYamamoto,M.		
FEATURES	FEATURES		
SOURCE	Source		
1. 1988	Location/Qualifiers		
/organism="Oryza sativa (japonica cultivar-group)"			
/mol_type="mRNA"			
/cultiyar="Nipponbare"			
/db_xref="taxon:39947"			
/clone="J013135D01"			

Alignment Scores:					
Pred. No.:	6,39e-13	Length:	1988		
Score:	497.00	Matches:	132		
Percent Similarity:	50.46%	Conservative:	33		
Best Local Similarity:	40.37%	Mismatch:	99		
Query Match:	30.29%	Indels:	64		
DG:	8	Gaps:	12		
US-10-630-636-7 (1-306) x AKO68138 (1-1988)					
OY	10	ThrThraaaalaaalaaalaaalAATPrThrrArgGuaspeAsLySaIaphEgluaSnAla	29		
Dd	490	ACCGGGGGGGGGGGGGGGCGGA-GCGTGTCTGCAGAGAGCAAGAAGTGTTTCAGCAGCGC	548		
OY	30	LeuLaalaCybaIlaaIaProProProAlaaspGlyAlaIProaPsaPsTrPhe	49		
Dd	549	CTGGGGGGCTCCGACC GGAGACCG---GAA GAGTGG---	584		
OY	50	AlaIaLeuLaalaaIaSerValProGlyLaarGerSaIaGlugluValaArGHisTrY	69		
Dd	585	GAGCGGGGGGGCTGCTCTGCCGGC--AAGACGTGGCCGACGTATAGCACCATAC	641		
OY	70	GUuLaLeuVaIGluAspVaIaIaIaIleAspaIagiYaRvaIProLeuProArGTrY	89		
Dd	642	GACGACCTCGAGAACGACGTCTGTACTACGAGCGCGGCTCTCTCCCTTCCCACATAC	701		
OY	90	AlaGlyGuIuSerAlaIaProProAspGlyAlaIaIaIaIaIaIaIaSerLys	109		
Dd	702	-----GGGCGCGCGCGCGCGCGCGCGCTCGGGTTCACCTTGACTGG	746		
OY	110	AspGlyGIHIArGArgAspGLuaRgrysglyGlyglYglYglYTYrAsp-----	126		
Dd	747	GACGGCGGC-----GATGACCCCCCGGCTTGGGTTCAAGCCTCCTGCTACATG	797		
OY	127	--glycilyySerySerySerySaIaGlugluNGuarGArLySgLYleProTrThr	145		
Dd	798	GTCGCGGCAAGCGCGCGCGCGCCCTGACCAGAGGCCAAGAAAGGGGTGCGCGTAGCA	857		
OY	146	GUlUGluGUHIsARgLeuPheluLeuGlyLeuAspLySPHeGlyLVyGlyASPTPRaG	165		
Dd	858	GAGGAGGAGCAAGATTGTTCTTGAATGGGGCTGAAGAAAGTACGGAGAGGTACGTGAGG	917		
OY	166	SerIleSeRaRGaSpHeVallIeserArgThrProTrGlnValaISerHiSaIaGln	185		
Dd	918	AACATCTCCCACTTCTGTACAGAGCGGACGCCAGACAGTGGCCACGACCGGAG	977		
OY	186	LysTyRPhellleaRguLeuAsnSerMetasnRgaPaRgaRgaRgeSerIleHisAp	205		
Dd	978	AAGTACTTCATCCGGCTCACTCCGGCGGCAAGACAGCCCGCTCAGATCCACGAC	1037		
OY	206	IleThrsErValThraIaglyAspGlnValaIaIaIngInGlyAlaProIIehrcly	225		
Dd	1038	ATCCACCAACCGTCAACCTCCCGCACAC-----GAC 1067			
OY	226	HieGlnaIaTrngLYaSpProAlaIaIaIaleuGlyProProGlyMeLyehisHis	245		
Dd	1068	CAC-----GGCAACCTCGCGCGCGCG-----CCGCGTCCGTCTACCGCG	1112		
OY	246	HisHisHisIspRoGly-----GlyAla	253		
Dd	1113	CACTCCTCTCTCGCGCGCGCGGTCTCGAGCAGTTCGGCGTCTGTGACGGCAAG	1172		
OY	254	ProProPrometPrometTySerAlaIaPro-----MetGlyHisProval	269		
Dd	1173	CTCTCGCGCGCGCGCTGCGCGCGCGCGGACACCACTTATGCCCCATCTTAC	1232		
OY	270	AlaGlyHisMetValProAla-AlaValaGlyThrProValaIaPheProProGlyHisAl	289		
Dd	1233	GCCCAAGTAAGATGAGAGCGGCAATTGCGACATCGCGCGCGCGCGCGCTGACGAC	1292		
OY	289	aProTyR-----ValaIaProValaGlyTYrProAl	299		

```

Db      1293  TCGGTTGTGGATGCAGATGCATGGGCGGACCTGTATGCAGCTCTCGGTTGACCAACCTGT 1352
Oy      299  aPpRoCiaLalYmEchE 305
          |||
          |||
          |||
Db      1353  TCCTGGAGCAGCAACCCAC 1371

RESULT 15
AB083027 1191 bp  mRNA  linear  PLN 12-MAY-2004
LOCUS     Glycine max mRNA for syringolide-induced protein 1-3-1a, complete
DEFINITION
cds.
ACCESSION AB083027
VERSION   AB083027.1  GI:119911576
KEYWORDS
SOURCE    Glycine max (soybean)
           Glycine max
           Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
           Glycine.
REFERENCE 1
AUTHORS   Hagihara,T., Hashi,M., Takeuchi,Y. and Yamaoka,N.
TITLE      Cloning of soybean genes induced during hypersensitive cell death
           caused by syringolide elicitor
JOURNAL    Planta 218 (4), 606-614 (2004)
PUBMED    14586656
REFERENCE 2 (bases 1 to 1191)
AUTHORS   Hagihara,T., Hashi,M., Takeuchi,Y. and Yamaoka,N.
TITLE      Direct Submission
JOURNAL    Submitted (28-MAR-2002) Naoe Yamaoka, Hokkaido University,
           Division of Biological Sciences, Graduate School of Science;
           Kitai0, Nishi8, Sapporo, Hokkaido 060-0810, Japan
           (E-mail:yamaoka@sci.hokudai.ac.jp, Tel:11-706-3482,
           Fax:11-706-4851)

FEATURES
     source          location/Qualifiers
     source          1..1191
                     /organism="Glycine max"
                     /mol_type="mRNA"
                     /cultivar="Hatosoy"
                     /db_xref="taxon:3847"
                     /clone="1-3-1-3"
                     /cell_type="suspension cultured cell"
                     /clone_1lb="lambda syr-Hatosoy"
                     /note="cDNA library constructed from cells created with
                     syringolide"
     gene            1..1191
                     /gene="1-3-1a"
     cds             155..856
                     /gene="1-3-1a"
                     /note="similar to Arabidopsis thaliana putative I-box
                     binding factor protein"
                     /codon_start=1
                     /product="syringolide-induced protein 1-3-1a"
                     /protein_id="BAB86892.1"
                     /db_xref="GI:19911577"
                     /translation="MFQESVSTRPQWTRHYDKLPERALLVVEEDLPDRMEKTIADQV
                     PGSKAVEREHYALVHDVEIDSGRVESYVDSDVAMPSCGAGISTWDNAGISF
                     GSGKQGENERKKGGPMTSEERHFLFLIGSKRGKQDMSISRNVVTRPTOVASHA
                     OKXFLRNSVTKERKRSISHDITTVDSNSHPMIDQTVPPPGSGIQSGQYSSNMH
                     DQMGTRISNYGFDN"

ORIGIN
Alignment Scores:
Pred. No.:      6 71e-13      length:      1191
Score:          492.00      Matches:      113
Percent Similarity: 56.40%      Conservative: 118
Best Local Similarity: 45.20%      Mismatches: 61
Query Match:    29.98%      Indels:      48
DB:             8            Gaps:      8
US-10-630-636-7 (1-306) x AB083027 (1-1191)
Oy      18  TrpHnRgSluNsPaSbPySaIaPheGluASnAlaLeuAlaIaCysAlaIaIaProPro 37

```

```

Db      194  TGAAGTCTGCTACCAAGACAGACTATTTCAGAGCGGCGCTTCGTGC----- 238
QY      38   ProAlaAspGlyGlyAlaProAsp-----AspTrpPheAlaAlaLeuAlaAsp 55
Db      239  -----GTCCGAGAGATCTCCCGGACCGGTGGAGAGATGCCGACGAC 283
QY      56   ValProGlyAlaArgSerAlaGluGluValArgArgHisTyrGluAlaLeuValGluAsp 75
Db      284  GTTCCCGGC---AAGTCGGCGGTGAGGTAAAGGAACACTACGAGCGCTCGTCCACGAC 340
QY      76   ValAlaAlaIleAspAlaGlyArgValProLeuProArgTyrAlaGlyGluGluSerAla 95
Db      341  GTCTTCGAGATCGACTCCGCGCGAGTCGAGGTCCCGAGTTACGTGATGACTCCGCTGCGC 400
QY      96   AlaProPheAspGlyAlaGlyAlaAlaAlaAlaSerLysAspGlyGlyHisArgArg 115
Db      401  ATGCGCGCC---AGTGGCGGTGCTGGAATTCTACGTGGACACCGCTAATCAGATC 454
QY      116  AspGluArgLysGlyGlyGlyGlyTyrAspGlyGlyLysSerCysSerLysAlaGlu 135
Db      455  TCA-----TTCCGTCCAGCTTAACAAACAAGCGCAG 487
QY      136  GlnGluArgArgLysGlyIleProTrpThrGluGluGluHisArgLysPheLeuGlu 155
Db      488  AACGAAAGGAAGAGGCGACCGCTTGACGACGAAACACAGCGCTTCTCATTTGGG 547
QY      156  LeuAspLysPheGlyLysGlyAspTrpArgSerLysSerArgAsnPheValIleSerArg 175
Db      548  CTTAGTAGTTGGAGAGGAGATTGGAGAACATTTCAAGAAATGTTGTGTGCAAGA 607
QY      176  ThrProThrGluValAlaSerHisAlaGlnLysTyrPheIleArgLeuAsnSerMetAsn 195
Db      608  ACACCAACTCAAGTAGTACGCAATGCCAGAAATCTTCTTCGCCAGAAATTCGGTGAAG 667
QY      196  ArgAspArgArgArgSerSerIleHisAspIleThrSerValThrAlaGlyAspGluVal 215
Db      668  AAAGAGAGAGAGATCAAGCATTCATGACATTAACGACTGTGACAGT----- 715
QY      216  AlaAlaGlnGlnGlyAlaProIleThrGlyHisGlnAlaThrGlyAsnProAlaAla 235
Db      716  -----AACTCAGCGCCTATG-----CCCATTTGATCAG 742
QY      236  AlaLeuGlyProProGlyMetLysHisHisHisHisHisProGlyGlyAlaProPro 255
Db      743  ACTTGGGTTCCTCT-----CCGGTGGTTCACCTGACG 775
QY      256  ProMetProMetTyrSerAlaAlaProMet 265
Db      776  CAGTCACAGCAATACCCATCAAGCAACATG 805

```

Search completed: July 30, 2005, 00:26:00
 Job time : 5168 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 29, 2005, 19:31:08 ; Search time 614 Seconds
(without alignments)
2950.229 Million cell updates/sec

Title: US-10-630-636-7

Sequence: 1 MTSQAATTTTAAATAAATWTR.....GHAPVVPVGYAPAPAKKHQ 306

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlh
-O=/cgn2_1/USPFO_epoc1/US10630636/runat_27072005_123226_12158/app.query.fasta_1.455
-DB=N Geneseq_16Dec04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOFCI=0
-LOOPEXT=0 -UNITs=bits -START=1 -END=1 -MATRIX=blowum62 -TRANS=human40.cdi
-LIST=45 -DOCCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10630636 @CGN_1.1.644 @runat_27072005_123226_12158 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC
-DEV TIMEOUT=150 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1:	Geneseq_16Dec04:*
2:	Geneseqn1980s:*
3:	Geneseqn190s:*
4:	Geneseqn2000s:*
5:	Geneseqn2001as:*
6:	Geneseqn2002as:*
7:	Geneseqn2002bs:*
8:	Geneseqn2003as:*
9:	Geneseqn2003bs:*
10:	Geneseqn2003cs:*
11:	Geneseqn2003ds:*
12:	Geneseqn2004as:*
13:	Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description
1	1641	100.0	918	12	AD000780	Rice Myb	Ad000780 Rice Myb
2	1641	100.0	1330	12	AD000777	Rice Myb	Ad000777 Rice Myb
3	678.5	41.3	1161	3	AA034213	Arabidops	AA034213 Arabidops
4	624.5	38.1	1432	10	AD046636	Thalecres	AD046636 Thalecres
5	624.5	38.1	1432	10	AD055829	Thalecres	AD055829 Thalecres

6	624.5	38.1	1432	12	AD161458	Ad161458 cDNA enco
7	624.5	38.1	1432	12	AD002450	Ad002450 Thalecres
8	457.5	27.9	867	6	ABK65275	ABK65275 Arabidops
9	457.5	27.9	867	10	AD030621	Ad030621 Plant yie
10	457.5	27.9	867	10	AD031552	Ad031552 Plant yie
11	457.5	27.9	867	12	AD144216	Ad144216 Plant tra
12	456	27.8	668	6	ABK65273	ABK65273 Arabidops
13	456	27.8	668	10	AD030307	Ad030307 Plant yie
14	456	27.8	668	12	AD144214	Ad144214 Plant tra
15	456	27.8	668	12	AA046672	AA046672 Arabidops
16	455.5	27.8	608	13	ACN45747	ACN45747 Cotton pr
17	445	27.1	891	3	AA039215	AA039215 Arabidops
18	428.5	26.1	866	6	ABK65370	ABK65370 Arabidops
19	428.5	26.1	866	10	AD030950	Ad030950 Plant yie
20	428.5	26.1	866	10	AD037148	Ad037148 Plant yie
21	428.5	26.1	866	12	AD141988	Ad141988 Plant tra
22	428.5	26.1	866	12	AD003516	Ad003516 Thalecres
23	428.5	26.1	866	3	AA047819	AA047819 Arabidops
24	423	25.8	804	10	AD029992	Ad029992 Plant yie
25	423	25.8	804	12	AD144406	Ad144406 Plant tra
26	420.5	25.6	896	6	ABK65271	ABK65271 Arabidops
27	420.5	25.6	896	10	AD046650	Ad046650 Thalecres
28	420.5	25.6	896	10	AD030965	Ad030965 Plant yie
29	420.5	25.6	896	12	AD141842	Ad141842 Plant tra
30	420.5	25.6	896	12	AD002514	Ad002514 Thalecres
31	419.5	25.6	574	3	AA056812	AA056812 Eucalyptu
32	412.5	25.1	574	12	AD093427	Ad093427 Cotton ex
33	390	23.8	461	12	AA056762	AA056762 Eucalyptu
34	386.5	23.6	576	13	ACN49160	ACN49160 Cotton pr
35	381.5	23.2	536	12	ADP91116	AdP91116 Cotton ex
36	368.5	22.5	580	13	ACN48912	ACN48912 Cotton pr
37	367	22.4	559	13	ACN49230	ACN49230 Cotton pr
38	363.5	22.2	559	12	AD000779	Ad000779 Rice Myb
39	351.5	21.4	954	12	AD000782	Ad000782 Rice Myb
40	344	21.0	954	3	AA057102	AA057102 Ficus rad
41	336.5	20.5	332	3	AA046782	AA046782 Arabidops
42	335.5	20.4	1354	3	AA031814	AA031814 DNA encod
43	335.5	20.4	1386	10	AD002038	Ad002038 Thalecres
44	335.5	20.4	1386	12	AD002038	Ad002038 Thalecres
45	333.5	20.3	349	3	AA056698	AA056698 Eucalyptu

ALIGNMENTS

RESULT 1	
AD000780	standard; cDNA, 918 BP.
ID	AD000780 standard; cDNA, 918 BP.
XX	
AC	AD000780;
XX	
DT	26-AUG-2004 (first entry)
XX	
DE	Rice Myb protein OSMYBS1 cDNA.
XX	
KW	ss; gene; OSMYBS1; rice; plant; transgenic; Myb; transcription factor.
XX	
OS	Oryza sativa.
XX	
FH	Key
FT	Location/Qualifiers
FT	1..918
FT	/*tag= a
FT	/partial
FT	/product= "OSMYBS1"
FT	/note= "No stop codon given"
XX	
PN	US2004107456-A1.
XX	
PD	03-JUN-2004.
XX	
PF	30-JUL-2003; 2003US-00630636.
XX	
PR	31-JUL-2002; 2002US-03999999P.
XX	

PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136382P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139482P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.

PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148655P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149930P.
PR 23-AUG-1999; 99US-0150566P.
PR 25-AUG-1999; 99US-0150567P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156569P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157855P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158212P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 18-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.

Dd				:::::	615
Oy	556	TACTCATCCGCAACTTTCGGCGGCAGACAAAGACGACGACATTCACAGACTA			
Dd	207	ThisServatThraInlaagIapInValaIaaIagin-----GIngIyAlaProIleThr			224
Oy	616	ACCACCGTAATCTCGAAGAGGAGCGCTTTGGAGACAACATAAGAGCTCATTGTGT			675
Dd	225	GlyHiaBglN 227			
Oy	676	GGAGATCAG 684			
RESULT 9					
ADDJ0621					
ID	ADDJ0621	standard; cDNA; 867 BP.			
XX	ADDJ0621;				
XX	AC				
XX	AD				
XX	15-JAN-2004	(first entry)			
DE		Plant yield-related polynucleotide clone G1641.			
XX		ds; transcription factor; transgenic plant; growth rate; senescence;			
XX		seed germination rate; plant vigor; seedling vigor.			
XX	OS	Arabidopsis thaliana.			
XX	WO2003013227-A2.				
XX	PN				
PD	20-FEB-2003.				
XX					
PF	09-AUG-2002;	2002MO-US025805.			
XX					
PR	09-AUG-2001;	2001US-0310847P.			
PR	19-NOV-2001;	2001US-0336049P.			
PR	11-DEC-2001;	2001US-0338692P.			
PR	14-JUN-2002;	2002US-00171466.			
XX					
XX	(MENDEL-) MENDEL BIOTECHNOLOGY INC.				
PI	Ratlcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JB;				
PI	Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Plineda O, Yu G;				
PI	Brown PE;				
DR	WPI; 2003-248221/24.				
DR	P-PADB; ADDJ0622.				
PT		New plant transcription factor polynucleotides and polypeptides, useful			
PT		in producing transgenic plants with commercially valuable properties, useful			
PT		as an alteration in a plant growth characteristic, e.g. growth rate			
PT		or apomixis.			
PS		Disclosure; SEQ ID NO 650; 454bp; English.			
XX					
CC	The invention relates to a number of isolated Arabidopsis thaliana cDNA				
CC	sequences and their encoded proteins which are especially transcription				
CC	factor related cDNAs and proteins. The isolated or recombinant plant				
CC	transcription factor polynucleotides and polypeptides are useful in				
CC	producing transgenic plants with commercially valuable properties, i.e.				
CC	modified or altered desirable traits as compared to a reference plant,				
CC	such as an alteration in a plant growth characteristic, e.g. growth rate,				
CC	germination rate of seeds, vigor of plants and seedlings, or leaf and				
CC	flower senescence. Sequence information related to the polynucleotides				
CC	and polypeptides can also be used in bioinformatic search methods. The				
CC	transgenic plant is useful for growing a progeny plant from a parent				
CC	plant. This sequence represents one of the cDNAs of the invention.				
XX					
SQ	Sequence 867 BP; 243 A; 189 C; 251 G; 184 T; 0 U; 0 Other;				
Alignment Scores:					
Pred. No.:	7,33e-18	Length:	867		
Score:	457.50	Matches:	103		
Percent Similarity:	58.02%	Conservative:	38		
Best Local Similarity:	42.39%	Mismatches:	55		

FH	Key	Location/Qualifiers
FT	CDS	1..867
FT		/tag= a
FT		/product= "transcription factor"
XX		
PN	WO2003013228-A2.	
XX		
PD	20-FEB-2003.	
XX		
PF	09-AUG-2002; 2002MO-US025808.	
XX		
PR	09-AUG-2001; 2001US-0310847P.	
PR	19-NOV-2001; 2001US-0336049P.	
PR	11-DEC-2001; 2001US-0338692P.	
XX	14-JUN-2002; 2002US-00171468.	
PA	(MEND-) MENDEL BIOTECHNOLOGY INC.	
PI	Heard JE, Riechmann JL, Creelman RA, Keddie J, Pilgrim ML;	
PI	Dubell AN, Jiang C, Ratcliffe O, Pineda O, Yu G, Brown PE;	
XX		
DR	WPI; 2003-248222/24.	
XX	P-PsDB; ADE31553.	
PT	New plant transcription factor polynucleotides and polypeptides, useful	
PT	in producing transgenic plants with commercially valuable properties,	
PT	i.e. modified desirable traits, e.g. salt stress resistance or tolerance	
PT	to freezing.	
XX		
PS	Disclosure; SEQ ID NO 119; 31pp; English.	
XX		
CC	The invention relates to a number of isolated cDNA sequences and their	
CC	encoded proteins which are especially transcription factor related cDNA s	
CC	and proteins. The isolated or recombinant plant transcription factor	
CC	polynucleotides and polypeptides are useful in producing transgenic	
CC	plants with commercially valuable properties, i.e. modified or altered	
CC	desirable traits as compared to a reference plant, e.g. salt stress	
CC	resistance, osmotic stress resistance, tolerance to freezing, drought,	
CC	low humidity tolerance, or radiation resistance. Sequence information	
CC	related to the polynucleotides and polypeptides can also be used in	
CC	bioinformatic search methods. The transgenic plant is useful for growing	
CC	a progeny plant from a parent plant. This sequence represents one of the	
CC	cDNAs of the invention	
XX		
SQ	Sequence 867 BP; 243 A; 189 C; 251 G; 184 T; 0 U; 0 Other;	
	Alignment Scores:	
	Pred. No.: 7..33e-18 Length: 867	
	Score: 457.50 Matches: 103	
	Percent Similarity: 58.02% Conservativeness: 38	
	Best Local Similarity: 42.39% Mismatch: 55	
	Query Match: 27.88% Indels: 47	
	DB: 10 Gaps: 9	
	US-10-630-636-7 (1-306) x ADE31552 (1-867)	
OY	1 MetTThSerGlnAlaAlaThrTrhTrhTrhAlaAlaAlaAlaAlaATPrThArg 20	
	:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	49 CTCATCGAGGAAACTTAAGACGGCGCTGCAGCTTCGTGAAGGTGCCACGTGACGGC 108	
OY	21 GluAspAPLyLeAlahegluaenAlaleuAlaAlaCySaAlaAlaProProAlaasp 40	
	:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	109 GCAGAACAAACAGCATTCGAAATGCTTTGGCGGTT-----TAGCAC 150	
OY	41 GlyGlyAlaProAspaAspArprpheAlaAlaIleuAlaAlaservalProGlyAlaarg 60	
	:::: :::: :::: :::: :::: :::: :::: ::::	
Db	151 GACAACATCTCTATCG-----TGCCAGAAGGTGCTCGGTATCCGGG--AAG 201	
OY	61 SerAlaGlugluValArgArgHisTyrglualAleuValGluIuspsValalaAlaileasp 80	
	:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	202 ACAAGTAGAGCACTAATTAGACAGTATAACGATTTCGAACCTGATGTCACACAGATCGAG 261	
OY	81 AlaGlyArvalProleuproArgTyralaglyGluInuSeralAlaProPro----- 98	

Dd		262 GCCGGTTTAAATCCCGGTCCCGGTTACATC-----ACCTGCCCGCTTTACT	309
Oy	--AspGlyValGlyAlaAlaAlaAlaSerLysAspGlyHisArgArgAspGlu	117	
Dd	310 CTAGATTGGGC-----	321	
Oy	118 ArgLysGlyValGlyValGlyTyrAspGlyLysSer-----CysSerLys	133	
Dd	322 -----GGCGCGGTGGCGGATGTAAcgggtttTAACCgGTCATCAgTTTGATMAA	375	
Oy	134 AlaGluGln-----GluArgArgLysGlyIleProTPThGln	146	
Dd	376 CGGTCCGAAGCCCGGTAGATCCCGGAGCTGAGCGGAAGAAGAGCCGTTCTTGACGAG	435	
Oy	147 GluGluHisArgLeuPheLeuLeuGluLysLeuAspLysPheGlyLysGlyAspTPArgSer	166	
Dd	436 GAAGAACACACAGCTATTTCATAATGGGTTTGAAGAAATATGAGAAAGAGATTGGAGAAAC	495	
Oy	167 HisSerArgAsnPheValHisSerArgThrProThrGlnValHisSerHisAlaGlnLys	186	
Dd	496 ATATCTCCGAACTTTGTGTATACGGGAACGGCAACACAAAGTAGCTACCCAGCCCAAAG	555	
Oy	187 TyrPheIleArgLeuAsnSerMetAsnArgAspArgArgSerSerIleHisAspIle	206	
Dd	556 TACTTATCTCCGCACTTCCGGCGCGCAAGACAGACAGACAGACATTCACGACATA	615	
Oy	207 ThrSerValThrAlaGlyAspGlnValAlaAlaGln-----GlnGlyAlaProIleThr	224	
Dd	616 ACCACCGTAAATCTCGAAGAGAGCGCTCTTGGAGACAATAAGACTCCATGTTGTT	675	
Oy	225 GlyHisGln	227	
Dd	676 GGAGATCAG	684	
RESULT 11			
ID	ADI44216	standard; DNA; 867 BP.	
XX	ADI44216;		
DT	22-Apr-2004	(first entry)	
DE	Plant transcription factor related polynucleotide #1697.		
XX			
KW	transgenic; plant; enhanced tolerance to abiotic stress;		
KM	glycosphate tolerance; hormone sensitivity; disease resistance;		
KM	sugar sensing; flowering; flower structure; stem bifurcation;		
KM	branching pattern; apical dominance; trichome; stem morphology;		
KM	root growth; root hair; seed development; cell proliferation;		
KM	cell differentiation; premature senescence; necrosis; plant size;		
KM	leaf morphology; seed morphology; seed biochemistry; root anthocyanin;		
KM	plant anthocyanin; light response; shade avoidance; bioinformatic;		
KM	transcription factor; de.		
XX	Unidentified.		
OS			
XX	US2004019927-A1.		
PN			
PD	29-JAN-2004.		
XX			
XX	25-FEB-2003; 2003US-00374780.		
PF			
XX			
PR	18-APR-2001; 2001US-00837944.		
XX			
PA	(SHER/) SHERMAN B. K.		
PA	(RIEC/) RIECHMANN J L.		
PA	(JIAN/) JIANG C.		
PA	(HEAR/) HEARD J E.		
PA	(HAAK/) HAAKE V.		
PA	(GREE/) CREELMAN R A.		
PA	(RATC/) RATCLIFFE O.		
PA	(ADAM/) ADAM L J.		


```
QY 107 AlAserLyAspGlyGlyHisArgArgAspGluArgLySGlyGlyGlyGlyTyrAsp 126
    |||||
Db 390 GGGCTTAACATGGC----- 404
QY 127 GlyGlyLySerCySerLyAlaGluGluArgArgLySGlyLeuProTyrThrGlu 146
    |||||
Db 405 -----GAGAGTGAACGCAAAAGAGAACTCCTTGACACAG 440
QY 147 GluGluHisArgLeuPheLeuGlyLeuAspLySPheGlyLySGlyAspTyrArgSer 166
    |||||
Db 441 AACGACACAAATTGTTCTGATCGATTAAAGATATATGTTAAGGAGATTGAGAGAGT 500
QY 167 IleSerArgAnPheValIleSerArgThrProThrGlnValAlaSerHisAlaGlnLys 186
    |||||
Db 501 ATCTCGAGAAACGTTGTGTGACGAGACACCGCAAGTCGCGAGTCAAGTCAGAG 560
QY 187 TyrPheIleArgLeuAsnSerMetAsnArgAspArgArgSerIleHisAspIle 206
    |||||
Db 561 TATTTCTGACACAGAACTCGGTGAGAGAGAGAGAAAGGTCGAGCATTCATGATATA 620
QY 207 ThrSerValThrAla-----GlyAspGlnVal-----AlaAlaGlnGln 219
    |||||
Db 621 ACTACGGTTGATGCTTGTGCTATGCGCTGCTTAACATGAGACTGAGTGGCCAAAC 680
QY 220 GlyAlaProIleThrGlyHisGln 227
    |||||
Db 681 GGGAGTCCTCTTCAGGCGCCGACG 704
```

Search completed: July 29, 2005, 22:59:56
Job time : 625 secs


```
QY 34 AlAlaIaProProAlaAspGlyGlyAlaProAspAspPheAlaAlaLeuAla 53
    |||
Db 377 TTCAGAGCTTCCCGC-----GATTGCTGGAGAAAGTACCG 342
QY 54 AlAspValProGlyAlaArgSerAlaGluValArgAryHsTyrGluAlaLeuVal 73
    |||
Db 341 TCGACGTCGCGCGC---AAAACCTAGAGAGATTAGCTCCACTACGAGACCTGGTG 285
QY 74 GluAspValAlaAlaIleAspAlaGlyArgValProLeuProArgTyrAlaGlyGlu 93
    |||
Db 284 GAAGATCTCAATAGATAGAGCGTCGCGCTGCTGCGCGTCTAC----- 237
QY 94 SerAlaAlaProProAspGlyAlaGlyAlaAlaAlaAspSerLysAspGlyGlyHis 113
    |||
Db 236 AGTTCCGCTTCTCGATGGC-----TCCGCAAGTAACTTAGGGGAG 195
QY 114 ArgArgAspGluArgLysGlyGlyGlyTyrAsp-----GlyGlyLysSer 130
    |||
Db 194 GAGGAGGCAAGAAAGGGGCGCGGCGGCGGCGGCAACCAAGGAGGAGAGGC 135
QY 131 CysSerLysAlaGluGluGluArgArgLysGlyTyrProTyrThrGluGluHisArg 150
    |||
Db 134 ---TCGAGTCCGATCAGAGAGAGAGAGGAGTTCCTTGCACCGAAGAAACAGG 78
QY 151 LeuPheLeuGlyLeuAspLysPheGlyLysGlyAspTyrAspSerLysSerAsn 170
    |||
Db 77 TTAATTTCTTAGGCTTAGATTAATATAGGAAAGCGATTGCGAATATTTCCCGAAC 18
QY 171 PheValIleSerArg 175
    |||
Db 17 TTGTGCTCACAAAG 3
```

```
RESULT 2
US-09-640-211A-1484
; Sequence 1484, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640.211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1484
; LENGTH: 335
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(335)
; OTHER INFORMATION: n = A,T,C or G
US-09-640-211A-1484
```

```
Alignment Scores:
Pred. No.: 2.18e-23 Length: 335
Score: 386.50 Matches: 78
Percent Similarity: 80.91% Conservative: 11
Best Local Similarity: 70.91% Mismatches: 8
Query Match: 23.55% Indels: 13
DB: 4 Gaps: 3
```

US-10-630-636-7 (1-306) x US-09-640-211A-1484 (1-335)

```
QY 116 AspGluArgLysGlyGlyGlyGly-----TyrAspGlyGlyLysSerCys 131
    |||
Db 7 GAGGAGGCAAGAAAGGGGCGCGGCGGCGGCGGCAACCAAGGAGGAGAAAGGC--- 63
QY 132 SerLysAlaGluGluGluArgArgLysGlyTyrProTyrThrGluGluHis----- 149
```

```
Db 64 TCGAGGTCGATCAGAGAGAAAGGGGATTGCTTGACCGAAGAGAAACAGCTGG 123
QY 150 -----ArgLeuPheLeuGlyLeuAspLysPheGlyLysGlyAsp 163
    |||
Db 124 GGTCTTCTTCTTGTCAATCGGTTATTCTTCTAGGCGCTAGATTAATATGGAAAGGAGAT 183
QY 164 TyrArgSerLysSerArgAspPheValIleSerArgThrProThrGluValAlaSerHis 183
    |||
Db 184 TGGCGAAGTATTTCCCGAACTTTGTTCGNCACAAGACCGCCTACGCAAGTTGCGAGCCAT 243
QY 184 AlaGluLysTyrPheIleAspLeuAspSerMetAsnArgAspArgArgAspSerTyr 203
    |||
Db 244 GCACAGAGATATTTTATCCGTCGAGCTCTGTAAACAAATAGAGAGCGATTCAGCAT 303
QY 204 HisAspIleThrSerValThrAlaGlyAsp 213
    |||
Db 304 CATGATTCTACTACCGTACGAGCACTGAGAGAC 323
```

```
RESULT 3
US-09-640-211A-1824
; Sequence 1824, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640.211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1824
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-1824
```

```
Alignment Scores:
Pred. No.: 2.68e-19 Length: 332
Score: 336.50 Matches: 72
Percent Similarity: 71.07% Conservative: 14
Best Local Similarity: 59.50% Mismatches: 23
Query Match: 20.51% Indels: 13
DB: 4 Gaps: 3
```

US-10-630-636-7 (1-306) x US-09-640-211A-1824 (1-332)

```
QY 64 GluValArgArgHisTyrGluAlaLeuValGluAspValAlaAlaIleAspAlaGlyArg 83
    |||
Db 4 GAGGTAGAGAGGCAATTACGAGCTTCTTGTGAGAGATGTGCTGTGATTGAGTCTGCGCGG 63
QY 84 ValProLeuProArgTyrAlaGlyGluGluSerAlaAlaProProAspGlyAlaGlyAla 103
    |||
Db 64 GTTGCTTGGCTGCTGATTTCT---GAAATTCGATACACCGCCGCAATGATGTGACA- 119
QY 104 AlAlaAlaAlaSerLysAspGlyGlyHisArgArgAspGluArgLysGlyGlyGly 123
    |||
Db 120 TCAAGTTGGGCGAATCT-----CACAAACA-----GCAGGC 149
QY 124 GlyTyrAspGlyGlyLys-SerCysSerLysAlaGluGluGluArgArgLysGlyTyr 143
    |||
Db 150 GGTTCCTGTAAAGCTCCCTCGGCAAGCATCGAAGCAGACGCAAAAGGCGCGCC 209
QY 143 OTTPTThrGluGluGluHisArgLeuPheLeuGlyLeuAspLysPheGlyLysGlyAs 163
    |||
Db 210 CTGAGCTGAGAGAGACAGACTCTTCTTGTGAGGATTAATTAATATGCAAAAGGTGA 269
QY 163 PTPArgSerLysSerArgAspPheValIleSerArgThrProThrGluValAlaSerHis 183
    |||
```



```
; SEQ ID NO 2038
; LENGTH: 1055
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-2038

Alignment Scores:
Pred. No.: 1,51e-16 Length: 1055
Score: 311.00 Matches: 70
Percent Similarity: 52.81% Conservative: 24
Best Local Similarity: 39.33% Mismatches: 42
Query Match: 18.95% Indels: 42
DB: 4 Gaps: 5

US-10-630-636-7 (1-306) x US-09-640-211A-2038 (1-1055)
QY 71 AAlaevAlaGluAAspValAlaAlaAlaAspAlaGluValProLeuProArg----- 88
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 299 GCTCTGCTCTCTCAATTGGGTTCTGTGACGAGCTCCATCAACAATGTCAAAGAAATTG 358
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 89 -----TyrAlaGluGluGluSerAlaAlaProProAspGly 100
Db 359 TTGATGATGTGTTCACACTGTGGCAGTGGCAGTGGCAGCTCCAGACCTGTCTGATAGA 418
QY 101 Ala-----GlyAlaAlaAlaAlaAlaSerLysAspGly-----GlyHis 113
Db 419 GGATCTGTCAAAATGTTGGGGTCAAGCTCATTTCTACAGACGATGGCATGGCCTGCATG 478
QY 114 ArgArgAspGluArgLysGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 125
Db 479 AGAAGAGACCTCAATATGGGCAATCTCGTCATTCACCTTACATTCATCAATCAC 538
QY 126 ---AspGlyGlyLysSerCysSerLysAlaGluGln----- 136
Db 539 TGTCTGGACAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 598
QY 137 -----GluArgGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 148
Db 599 GTTCATTCTCCAGCAATGCACGCCGAGAGAAAAAGCCGCCATGCTGCGAGGAAG 658
QY 149 HisArgLeuPheLeuLeuGlyLeuAspLysPheGlyLysGlyAspLysSerLysSer 168
Db 659 CACGAGATGTTCTTGTATGACCTGGAAGAGCTTGGAAGGTGCTGGAAGGATATCC 718
QY 169 ArgAsnPheValIleSerArgThrProThrGlnValAlaSerHisAlaGlnLysTyrPhe 188
Db 719 AGGAATTTTGTGACGACAGAACACCAACCAAGTACGACGACCAAGCAAGTATTTT 778

QY 189 IleArgLeuAsnSerMetAsnArgAspArgArgArgSerLysIleHisAspIle 206
Db 779 CTAAAGCAGAGCAATCTTAATMAAAGAAACGTCGATCCAGTCTTTGATATG 832

RESULT 11
US-09-640-211A-1960
; Sequence 1960, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1960
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1960
```

```
Alignment Scores:
Pred. No.: 1.83e-16 Length: 599
Score: 306.00 Matches: 73
Percent Similarity: 55.11% Conservative: 24
Best Local Similarity: 41.48% Mismatches: 27
Query Match: 18.65% Indels: 52
DB: 4 Gaps: 5

US-10-630-636-7 (1-306) x US-09-640-211A-1960 (1-599)
QY 55 SerValProGluValArg-SerAlaGluGluValArgArgHis-TyrGluAlaLeuValG 74
Db 130 GCCGTGCCGTGCGAGATGCTCCGCCACGCGCGCCAGTCCGCCAGCTGCCGCCGCG 189
QY 74 LuAspValAlaAlaAlaAspAlaGly----- 84
Db 190 ACAGAGACTCGCGCGCGCGCGCGCGCGCGCGAGATCATGCTTCCGGGTGCGGGTGG 249
QY 85 -----ProLeuProArgTyrAlaGlyGluG 93
Db 250 TCGTGACCTGATGAGGAACTGCGTGAAGCTTGAACAACCTGTCTCAATACAGACCCGC 309
QY 93 LuSerAlaAlaProProAspGlyAlaGlyValAlaAlaAlaAlaSerLysAspGlyGly 113
Db 310 AGAAGCGAATTCGCGCCACCGCCAGCGCGCGAGC-----GGCGCA 351
QY 113 isArgArgAspGluArgLysGly-----GlyG 122
Db 352 ACAAGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 411
QY 122 LysGlyGlyTyrAspGlyGlyLysSerCysSerLysAlaGluGlnGluArgLysGly 142
Db 412 GTGCG-----CCGAGCGCGAAGAGAGAG 435
QY 142 LepProThrGluGluGluHisArgLeuPheLeuGluLeuAspLysPheGlyLysG 162
Db 436 TTCTTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 495
QY 162 LysAspLysSerLysSerArgAsnPheValIleSerArgThrProThrGlnValAla 182
Db 496 GAGATTGAGAGAGATTCAGAGAACTTTGTGAAGACCGCAGCCGCTCAGAGTCCGA 555
QY 182 eHisAlaGlnLysTyrPheIleArgLeuAsnSerMetAsnArg 196
Db 556 GCCATGCCGAGAAATTTCTCGCGCGAGCAACTCAATCCG 599

RESULT 12
US-09-640-211A-1984
; Sequence 1984, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1984
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-1984

Alignment Scores:
Pred. No.: 1.35e-16 Length: 332
Score: 303.50 Matches: 56
Percent Similarity: 81.61% Conservative: 15
```



```
QY 166 SerIleSerArgAsnPheValIleSerArgThrProThrGlnValAlaSerHisAlaGln 185
    |||.....|||.....:..|||.....|||.....|||.....|||.....|||
Db 469 GGCATTTCGAGAGACTTTGTGAGAGCTCGACGCCAGCCAGCTCGAGCATGCCCAA 528
QY 186 LysIYrPheIleArgLeuAsnSerMetAsnArgAspArgArgSerSerIleHisAsp 205
    |||.....|||.....:..|||.....|||.....|||.....|||.....|||
Db 529 AAATACCTTCTCCGCCGAGACACCTTAATCG----- 561
QY 206 IleThrSerValThrAlaGlyAspGlnValAlaAlaGlnGlnGlyAlaProIleThrGly 225
    561 ----- 561
QY 226 HisGlnAlaThrGlyAsnProAlaAlaAlaLeuGlyProProGlyMetLysHis 245
    562 -----CCGTCGCCGCC-----GGACTTAGCCTTGTT 588
Db 246 HisHisHisPro 250
    |||.....|||.....|||.....|||.....|||.....|||
Db 589 TGACATCACCCCG 603
```

```
RESULT 15
US-09-640-211A-2059
; Sequence 2059, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2059
; LENGTH: 624
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-2059
```

```
Alignment Scores:
Pred. No.: 1.86e-15 Length: 624
Score: 294.00 Matches: 62
Percent Similarity: 67.92% Conservative: 10
Best Local Similarity: 58.49% Mismatches: 20
Query Match: 17.92% Indels: 14
DB: 4 Gaps: 3
```

US-10-630-636-7 (1-306) x US-09-640-211A-2059 (1-624)

```
QY 95 AlaAlaProProAspGlyAlaGlyAlaAlaAlaAlaSerLysAspGlyGlyHisArg 114
    |||.....|||.....:..|||.....|||.....|||.....|||.....|||
Db 336 GCTGACCCCTCCGAGCCGCCCTCTGCCCTCTCTCTGCCGCG----- 377
QY 115 ArgAspGluArgLysGlyGlyGlyGlyTy-----AspGly-----GlyLysSer 130
    378 -----GGCGGCGCGCCAGTGGCTATCTCTGTATGGTCTTGTGAAGCTCC 425
Db 131 CysSerLysAlaGlnGlnArgArgLysGlyLleProTPTHrGlnGlnGlnHisArg 150
    |||.....|||.....:..|||.....|||.....|||.....|||.....|||
Db 426 ACTTCCTCCCAATTCCTCGAGCGGAGAAAGGTGTGCTGCATGCACAGAGGAAACATAGA 485
QY 151 LeuPheLeuLeuGlyLeuAspLysPheGlyLysGlyAspTPAysSerIleSerArgAsn 170
    :..|||.....|||.....|||.....|||.....|||.....|||.....|||
Db 486 ATGTTTTCCTAGGTTTTCAGAAAGCTTGCAAGGTGATTCGAGAGATTCACGGAAT 545
QY 171 PheValIleSerArgThrProThrGlnValAlaSerHisAlaGlnLysTyPheIleArg 190
    |||.....|||.....:..|||.....|||.....|||.....|||.....|||
Db 546 TTGTGATTAACAGACAGACCTACACAGGTAGCCATGACAGAAATATTTATTATCGA 605
QY 191 LeuAsnSerMetLysArg 196
```

Db 606 CAGAGCAATATGACTAGA 623

Search completed: July 30, 2005, 01:28:20
Job time : 224 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 29, 2005, 22:49:43 ; Search time 716 Seconds

(without alignments)
2766.119 Million cell updates/sec

Title: US-10-630-636-7
Perfect score: 1641
Sequence: 1 MTSQNAATTTTAAAAAATW.....GHAPVVPVGPAPPAKMHQ 306

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgnr1/USFPO_epool/US10630636/runat_27072005_123229_12221/app_query.fasta.1.455
-DB=Published Applications_NA -OFMT=faetap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=10sum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=plo -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10630636@cgnr1.1.789@runat_27072005_123229_12221
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications_NA:*

- 1: /cgnr2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgnr2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgnr2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgnr2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgnr2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgnr2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 7: /cgnr2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 8: /cgnr2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgnr2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgnr2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 11: /cgnr2_6/ptodata/1/pubpna/US09C_NEW_PUB.seq:*
- 12: /cgnr2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgnr2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 14: /cgnr2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgnr2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgnr2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgnr2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgnr2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgnr2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
- 20: /cgnr2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
- 21: /cgnr2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
- 22: /cgnr2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 23: /cgnr2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
- 24: /cgnr2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 25: /cgnr2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 26: /cgnr2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1641	100.0	918	19 US-10-630-636-4	Sequence 4, Appl1
2	1641	100.0	1330	19 US-10-630-636-1	Sequence 59484, A
3	1641	100.0	1433	19 US-10-437-963-59484	Sequence 168067, A
4	1153.5	70.3	1299	20 US-10-425-115-168067	Sequence 2650, Ap
5	1087	66.2	1516	20 US-10-739-930-2650	Sequence 11058, A
6	727	44.3	1281	18 US-10-425-114-11058	Sequence 78758, A
7	727	44.3	1318	18 US-10-424-599-78758	Sequence 27578, A
8	707.5	43.1	1523	19 US-10-437-963-27578	Sequence 4609, Ap
9	698.5	42.6	823	18 US-10-425-114-4609	Sequence 57473, A
10	645	39.3	2366	18 US-10-424-599-57473	Sequence 7669, Ap
11	634	38.6	965	19 US-10-767-701-9031	Sequence 35, Appl
12	628	38.3	1230	18 US-10-425-114-7669	Sequence 157, App
13	624.5	38.1	1432	14 US-10-286-264-35	Sequence 853, App
14	624.5	38.1	1432	15 US-10-286-264-187	Sequence 157, App
15	624.5	38.1	1432	17 US-10-302-267-157	Sequence 157, App
16	624.5	38.1	1432	18 US-10-412-6998-863	Sequence 13945, A
17	571	34.8	800	20 US-10-425-115-113945	Sequence 52607, A
18	525.5	32.0	3091	20 US-10-425-115-113947	Sequence 183212, A
19	515.5	31.5	1896	20 US-10-425-115-184212	Sequence 34443, A
20	497.5	30.3	1140	19 US-10-437-963-34443	Sequence 52609, A
21	494	30.1	1613	20 US-10-425-115-52609	Sequence 2755, A
22	492	30.0	1202	18 US-10-425-114-29755	Sequence 34323, A
23	492	30.0	1363	18 US-10-424-599-34323	Sequence 52607, A
24	489	29.8	1416	18 US-10-424-599-34322	Sequence 440, App
25	488	29.7	1893	20 US-10-425-115-52607	Sequence 52610, A
26	484	29.5	2481	20 US-10-425-115-440	Sequence 96811, A
27	482	29.4	1388	20 US-10-425-115-52610	Sequence 7450, Ap
28	480	29.3	1349	19 US-10-437-963-96811	Sequence 8016, Ap
29	473	28.8	1021	18 US-10-425-114-7450	Sequence 183322, A
30	470	28.6	1039	18 US-10-424-599-8016	Sequence 183317, A
31	470	28.6	1793	20 US-10-425-115-183322	Sequence 253, App
32	469	28.6	1028	20 US-10-425-115-183317	Sequence 653, App
33	457.5	27.9	867	10 US-09-934-455-253	Sequence 2679, Ap
34	457.5	27.9	867	17 US-10-225-066A-653	Sequence 2679, Ap
35	457.5	27.9	867	17 US-10-225-067-119	Sequence 2679, Ap
36	457.5	27.9	867	17 US-10-374-780A-2679	Sequence 2679, Ap
37	457.5	27.9	867	22 US-10-225-066A-653	Sequence 2679, Ap
38	456	27.8	668	10 US-09-934-455-249	Sequence 2679, Ap
39	456	27.8	668	17 US-10-225-066A-339	Sequence 339, App
40	456	27.8	668	17 US-10-374-780A-2679	Sequence 3577, Ap
41	456	27.8	668	22 US-10-225-066A-339	Sequence 339, App
42	455.5	27.8	608	19 US-10-021-323-528	Sequence 528, App
43	443.5	27.0	1523	19 US-10-437-963-5655	Sequence 168064, A
44	441	26.9	548	20 US-10-425-115-168064	Sequence 7392, Ap
45	437	26.6	1136	18 US-10-425-114-7392	

ALIGNMENTS

RESULT 1
US-10-630-636-4
; Sequence 4, Application US/10630636
; Publication No. US20040107456A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Su-May
; TITLE OF INVENTION: PLANT MYB GENES
; FILE REFERENCE: 08919-088001
; CURRENT APPLICATION NUMBER: US/10/630, 636
; PRIOR FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US 60/399,999
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 918

QY	161	aserrhisalaglnlytyrrphei eargleuabnserrmet sena rgaspa rga g arg	200
Db	560	GGGAGCCACGCGGAGAGTACTTCATCCGCTCAACTCCATGAA CCGCGAC CCGCCCGC	619
QY	201	SerSer le hi sa p le thr Ser Val Thr Ala Gly Val Ala Ala Gln Gly	220
Db	620	TCCAGCATTCACGACATC TACCA GCGTC ACCGCCGGCATATAGTTCGCGCGCAGCAGGCG	679
QY	221	Ala Pro le Thr Gly His Gln Ala Thr Gly Val Pro Ala Ala Ala Gly Pro Pro	240
Db	680	GGCCCGATCAGCGGCGCACGAGCGCA CCGCGCA CCCGCGGCGGCGGTGCGCCGCGG	739
QY	241	Gly Met Val His His His His His Pro Gly Gly Ala Pro Pro Pro Met Pro Met Tyr	260
Db	740	GGCATGAA GAC CAC CAC CAC CAC CAC CCCGGGGGCGCGCGCGCCCATGCCCATGTAC	799
QY	261	Ser Ala Ala Pro Met Gly His Pro Val Ala Gly His Met Val Pro Ala Ala Val Ala Gly Thr	280
Db	800	AGCGCGCGCCCATGAGGCC CAC CCCGTCC CGGGCCACATGAGTCC CGCGCGCGGTGGGACG	859
QY	281	Pro Val Val Phe Pro Pro Gly His Ala Pro Tyr Val Val Pro Val Gly Tyr Pro Ala Pro	300
Db	860	CCGGTGGTGTTC CCCGCGGGCC ACGCGCGCTAGCTGTG CCGTCGAGCTAC CCGCGCGCT	919
QY	301	Pro Ala Val Met His Gln	306
Db	920	CCGGCCAA GATG CAC CAA	937
RESULT 3			
US-10-437-963-59484/c			
Sequence 59484, Application US/10437963			
Publication No. US20040123343A1			
GENERAL INFORMATION:			
APPLICANT: La Rosa, Thomas J.			
APPLICANT: Kovalic, David K.			
APPLICANT: Zhou, Yihua			
APPLICANT: Cao, Yongwei			
APPLICANT: Wu, Wei			
APPLICANT: Boukharov, Andrey A.			
APPLICANT: Barbazuk, Brad			
APPLICANT: Li, Ping			
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With			
FILE REFERENCE: 38-21(53221)B			
CURRENT APPLICATION NUMBER: US/10/437,963			
CURRENT FILING DATE: 2003-05-14			
NUMBER OF SEQ ID NOS: 204966			
SEQ ID NO 59484			
LENGTH: 1433			
TYPE: DNA			
ORGANISM: Oryza sativa			
FEATURE:			
OTHER INFORMATION: Clone ID: PAT_MRT4530_61102C.1			
US-10-437-963-59484			

Alignment Scores:		
Pred. No.:	1.59e-150	1433
Score:	1641.00	Matches: 306
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	19	Gaps: 0
US-10-630-636-7 (1-306) x US-10-437-963-59484 (1-1433)		
QY	1	MeetrhSerglnaiaaiaaThrThrThrThrThrAlaiaaiaaiaaiaaTrpThrArg 20
bB	1433	ATGACCTCCCAAGCGCGGCGAGCAGCAGCACCACGCGCGCGCGCGCGCGTGGACCAAG 1374
QY	21	GiuaappApPlyalaPheglubenaiaaleuaiaaiaCysaiaaiaaPpPpPpPpPpPpPp 40
bB	1373	GAGGACGACACAGCGCGTTCAGAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGAC 1314

QY	41	AllylGlyValAlaProAspAspSerTrpPheAlaIleuAlaIleValAlaSerValProGlyValArg	60
Db	1313	GAGGCGCGCCGAGACGACGACTGTTCCCGCGCTCCGCGCAGCGTGCCTCGGCGCGAAG	1254
QY	61	SerAlaGluGluValAlaArgArgHisTyrGlyLualaLeuValGluAspValAlaIleLeuAsp	80
Db	1253	TCGCGCGAGAGAGTTCGCGAGGCACCTACAGAGCGCTGCTGAGAGAGTTCGCGGCATTCAGC	1194
QY	81	AlaGlyValArgValProLeuProArgTyrAlaGlyGluGluSerAlaAlaProProAspGly	100
Db	1193	GCGGCGCCGCTCCGCTCCGCTCCGCTACCCCGGCGGAGGTCCGCGCGCCGCCACGGA	1134
QY	101	AlaGlyAlaIleAlaAlaAlaSerTyrAspGlyGlyHisArgArgTyrAspGluArgLysGly	120
Db	1133	GCCGAGCCGCGCGCGCGCTCCAGGACCGCGGACACCGCGCGAGCGGACGCGCAAGGC	1074
QY	121	GlyGlyGlyGlyTyrAspGlyGlyLysSerCysSerLysAlaGluGluGluArgLys	140
Db	1073	GCGCGCGCGGGGTACGACGCGCGGACAGAGTCTGCTCAAGCGGACGAGACGAGAGCGCAAG	1014
QY	141	GlyIlePcoTPrThrGluGluGluHisArgLeuPheLeuLeuGlyLeuAspLysPheGly	160
Db	1013	GGCATCCCATGACGAGAGAAAGACAGCGCTGTTCTTGCGGGCTGACAAAGTTCCGC	954
QY	161	LysGlyAspTrpArgSerLieserArgAspPheValIleSerArgThrProThrGluVal	180
Db	953	AAGGGGAGCTGGCGGAGCATCTCGCGCACTTCGTCATCTCGCGGACCGCAACGCAAGGTG	894
QY	181	AlaSerHisAlaGluLysTyrPheIleArgLeuAsnSerMetAsnArgAspArgArg	200
Db	893	GCGAGCCACGCGCAGAACTACTTCATCCGCTCAACTCATGAAACGCGACCGCGCGCCG	834
QY	201	SerSerIleHisAspIleThrSerAlaThrAlaGlyAspGluValAlaAlaGluGlnGly	220
Db	833	TTCACACATCCACGACATACACAGCGCTCACCGCGCGCATCAAGTTCGCGCGCAGCAGCGGC	774
QY	221	AlaProIleThrGlyHisGluAlaIleThrGlyAsnProAlaIleAlaIleAlaGluGlyProPro	240
Db	773	GCCCCGATCACCGGCGCACGAGCCACGCGGCAACCCCGCGCGCGCTGGGCCCGCG	714
QY	241	GlyMetLysHisHisHisHisHisIleProGlyGlyAlaProProProwMetProMetTyr	260
Db	713	GGCATGAAGCACACACACACACACACACCGCGCGCGCGCGCCCATGCTCCCATGTGAC	654
QY	261	SerAlaAlaProMetGlyHisProValAlaGlyHisMetValProAlaAlaValaGlyThr	280
Db	653	AGCGCGCGGCCCATAGGCGCACCCGCTCCCGCGGACACATGTCGCCCGCGCGCTGCGACG	594
QY	281	ProValValPheProProGlyHisHisAlaProTyrValValProValGlyTyrProAlaPro	300
Db	593	CCGGTGGGTTCGCGCGCGGCGCAGCGCGTACGTCGTGCGCGTGGCTAACCGCGCGCA	534
QY	301	ProAlaLysMetHisGln	306
Db	533	CCGCGCAAGATGCACCA	516

```

/ RESULT 4
/ US-10-425-115-168067
/ Sequence 168067, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovacic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants
/ FILE REFERENCE: 38-21 (53222) B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 168067

```



```
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 78776
; LENGTH: 1318
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: clone ID: PAT_MRT3847_4214C.1
US-10-424-599-78776

Alignment Scores:
Pred. No.: 2,656-61      Length: 1318
Score: 727.00           Matches: 164
Percent Similarity: 62.58%      Conservative: 40
Best Local Similarity: 50.31%    Mismatches: 70
Query Match: 44.30%           Indels: 52
DB: 18                   Gaps: 13

US-10-630-636-7 (1-306) x US-10-424-599-78776 (1-1318)

Qy 10 ThrThraAlaAlaAlaAlaAlaAlaTThraGluAspAlaPheGluAsnAla 29
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 72 ACCATGTCATCAAGTGGAAACCATTTGGAGCTATGAGGAAGAAAAGCATTTGAGAAATGCC 131

Qy 30 LeuAlaAlaCysAlaAlaProProAlaAspGlyAlaProAspAspTTP--- 48
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 132 ATACCTATG-----CATTGATT 149

Qy 49 -----PheAlaAlaLeuAlaAlaSerValProGlyAlaArgSer 61
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 150 GAGAAATCTCAAAAGAGCAATGGAGAAATTTGCTTACAGATTCC--AGCAAAAGC 206

Qy 62 AlaGluAluValaArgArgHisTyrGluAlaLeuValaGluAspValaAlaAlaIleAspAla 81
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 207 ATGGAAGATGTGAAGCAACATTACAGGTTCTAGTAGAGATGTAAAGTCAAATAGAGCA 266

Qy 82 GlyArgValProLeuProArgTyrAlaGlyGluGlu-----SerAlaAlaProProAsp 99
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 267 GGTCACTATCATCTCCCAAACTATGCTCTCTGAGAAATCACATCTTCAAAATAGAGAC 326

Qy 100 GlyAlaGlyAlaAlaAlaAlaAlaAlaSerLysAspGlyAlaHisArgAspGluArgLys 119
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 327 TTTCATGGCTCTTCCAGGCGCACAGCTCAGATAAAGATCAAAATTTGTAATTATGAGCGT 386

Qy 120 GlyGlyGlyGly---GlyTyrAsp-----GlyGlyLysSer 130
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 387 GGTTTCTGGGTAGGACATGACTCCACCATATGATGTGTAAAGAGGCGTTGCA 446

Qy 131 CysSerLysAlaGluGluGluArgArgGlyLysProTyrPThrGluGluHisArg 150
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 447 AGGTCA--TCGAACAGAAAGAAAGAAAGAAATCCCATGACTGAAGAGAAACACAGG 503

Qy 151 LeuPheLeuLeuGlyLeuAspLysPheGlyLysGlyAspTTPArgSerLysSerArgAsn 170
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 504 TTATTTTTCATCTGCTTAGAAAAGTTTGGAAAAGAGATTGGAAGAGATTTCAGAGGAC 563

Qy 171 PheValIleSerArgTyrProThrGluValaLaserHisAlaGluLysTyrPheIleArg 190
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 564 TTTGTGTATCTAGACACTCCACCTCACTGCAAGCCATGCACAAAAGTACTTCAATAGG 623

Qy 191 LeuAsnSerMetAsnArgAspArgArgArgSerSerLysIleAspIleThrSerValThr 210
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 624 TTGAATTCATGAATAGAGACAGAGAGGCTTATGATATCATGATATCATGATGATGAC 683

Qy 211 AlaGlyAspGluValaAlaAlaGluGluGlyAlaProIleThrGlyHisGluAlaThrGly 230
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 684 AATGAGAGT-----GTGGCTAATTAACCAACACACTATTATACAGGCGCAGCATAGAGCA 737

Qy 231 AsnProAlaAlaAlaAlaLeuGlyProProGlyMetLysHisHisHisHisPro 250
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 738 ATTCCTTCAACACAAATGGGTGTAGACATATCCCTTAAGCATTAAGTTCAAGGCTAC-- 794

Qy 251 GlyGlyAlaProProProMetProMetTyrSerAlaAlaProMetGlyHisProValAla 270
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
```

```
Db 795 -----ATACCACCTGCTTAGCAGCATAT---GGAACACACAGTTGACATCTGTGGCT 845

Qy 271 -----GlyHisMetValProAlaAlaValGlyThrProValValPheProProGly 287
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 846 GCTCTCTCAGGGCAGACATGACA--TTGACAGTTGGCAGCTCTGTCATGCTCTCCCTGGA 902

Qy 288 -----HisAlaProTyrValValProValGlyTyrProAlaPro 300
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 903 CCCCACCCCAGTCTATCTCTCATTCACCTTAATGTTCTCTCTGTTACCAATGCGCA 962

Qy 301 ProAlaLysMetHisGln 306
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 963 CTTCCAAACATGATCA 980

RESULT 8
US-10-437-963-27578
; Sequence 27578, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Mu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/437, 963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 27578
; LENGTH: 1523
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: clone ID: PAT_MRT4530_32258C.1
US-10-437-963-27578

Alignment Scores:
Pred. No.: 2,486-59      Length: 1523
Score: 707.50           Matches: 167
Percent Similarity: 58.57%      Conservative: 21
Best Local Similarity: 52.02%    Mismatches: 76
Query Match: 43.11%           Indels: 57
DB: 19                   Gaps: 13

US-10-630-636-7 (1-306) x US-10-437-963-27578 (1-1523)

Qy 12 AlaAlaAlaAlaAlaAlaAlaAlaTTPThraGluAspAlaPheGluAsnAlaLeuAla 31
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 466 TCCGTTGGCGGCGGTGGTGGAGCGCGCAGACAGAAAGCGTTGAGAAAGCGGTCGCG 525

Qy 32 AlaCysAlaAlaProProProAlaAspGlyAlaProAspAspAspTTPheAlaAla 51
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 526 ACGGTG-----GGGATGACGAGAGAGAGAGAGGAGCGGTTGTGGAGAG 570

Qy 52 LeuAlaAlaSerValProGlyAlaArgSerAlaGluGluValaArgArgHisTyrGluAla 71
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 571 CTACCGAGAGCGCGTGGAGGGG--AAGACGCGCAGACAGAGGTGAGCGGCACTACGAGCTG 627

Qy 72 LeuValGluAspValaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 91
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 628 CTGGTGAGAGCGTCGACGCGATCGAGCGCGGCGGCGGCTCTCGTGTAGCGCGGC 687

Qy 92 GluGluSerAlaAlaProProAspGlyAla-----GlyAlaAlaAlaAlaAlaSer 108
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 688 -----GACGGGGCGCTCGACGAGAGGCGCTCTGCGGAGGTGGAGAG 726

Qy 109 LysAspGlyGlyHisArgArgAspGluArgLysGlyGlyGlyGlyTyrAspGlyGly 128
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
```


DB: 18 Gaps: 15

US-10-630-636-7 (1-306) x US-10-424-599-57473 (1-2366)

QY 18 TrpThrArgGluAspAlaPheGluAsnAlaLeuAlaIleAlaIleAlaProPro 37

DB 836 TGGGAAAAAGAAACGATTAAGCATTTGGAAAATGCTTGGCAAT----- 880

QY 38 ProAlaAspGlyAlaProAspAsp-----AspTrpPheAlaAlaLeuAlaIleSer 55

DB 881 -----CACCTGAGATGCTTCAGACCGGTGGAGAGATGACAGTGCAT 925

QY 56 ValProGlyAlaAspSerAlaGluValArgArgHisTyrGlyAlaLeuValGluAsp 75

DB 926 GTACCAAGGG--AAACCTTGGAAGATTAAACCACTATGAGCTCTGTGTGAAGAT 982

QY 76 ValAlaAlaIleAspAlaGlyArgValProLeuProArgTyr-----AlaGly 91

DB 983 GTTAGCCAGATCGAATCCGGTTATGTGCTTTACCATCTTATATCTTCTCCAGAGGC 1042

QY 92 GluGluSerAlaAlaProProAspGlyAlaGlyAlaAlaAlaAlaSerLysAspGly 111

DB 1043 TCACAAGCCATGCTGATGAAAGAGAGCTGGC-----AAGAAGGA 1084

QY 112 GlyHisArgArgAspGluArgGlyGlyGlyGlyGlyGlyTyrAspGlyGlyLysSerCys 131

DB 1085 GGCACACTCTGGAAATAGTAATTAACGAATCT-----AATCATGGAACTTAAGCT 1132

QY 132 SerLysAlaGluGluGluArgArgGlyLysLeuProTrpThrGluGluGluHisArgLeu 151

DB 1133 TCAGATCAAGTCAAGACGAAACGAAAGGGTTTGCATGACAGACGATGAACACAGGTTA 1192

QY 152 PheLeuLeuGlyLysAspLysPheGlyLysGlyAspTrpArgSerLysSerArgAspPhe 171

DB 1193 TTCCTCTGGCTTGGAAGATAGGAAAGTACGCGGAGATATATCAAGAACTTT 1252

QY 172 ValIleSerArgTrpProThrGluValAlaSerHisAlaGluLysTyrPheIleArgLeu 191

DB 1253 GTGTGTGCAAGAACGCCCTACCAAGTAGCAAGCCATGCCAAAAATCTTCATTCGCTCG 1312

QY 192 AsnSerMetAsnAspArgArgArgSerSerIleHisAspIleThrSerValThrAla 211

DB 1313 AACTCGATGAATTAAGACAGAGCGATCCAGCATACAGATACACCAAGTGTAAACAT 1372

QY 212 GlyAspGlnValAlaAlaGlnGlnGlyAlaProIleThrGlyHisGlnAlaThrGly--- 230

DB 1373 GGAAGAT---GTTTCAGCAGCTCAAGGA---CCAATTACTGCT---CAACAACAGGTTCT 1423

QY 231 -----AsnProAlaAlaAlaLeuGlyProProGlyMetLysHisHisHisHis 248

DB 1424 GCGGATTAATCTGCTGGAAAATCAACCA----- 1456

QY 249 HisProGlyGlyAlaProProProMetPro-----MetTyrSer 261

DB 1457 -----GCCCAACAGCCCAACCTGCTGCTGGGTGAGTAAGAAATTATGCT 1501

QY 262 AlaAlaProMetGlyHisProValAlaGlyHisMetValProAlaAlaValAlaGlyThrPro 281

DB 1502 GGTCTTACCATTTGACAACTATAGAGAGACCCCTCGTA---TCTGAGTTGGACCCCA 1558

QY 282 ValValPhe-----ProProGlyHisAlaProTyr----- 291

DB 1559 GTGATGAACTTCTCTCCCGCAGCACATGTGATATGCTCGAGACACCGGTCTCTGGG 1618

QY 292 -----ValValProValGlyTyrProAlaPro 300

DB 1619 GCAAGTTGTTCTGGAGACCAATGAATTAAGTCTGTCATACCCATTAGCCA 1672

RESULT 11

US-10-767-701-9031/c

; Sequence 9031, Application US/10767701

; Publication No. US20040172684A1

; GENERAL INFORMATION:

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yina

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with

TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement

FILE REFERENCE: 38-21(5353)B

CURRENT APPLICATION NUMBER: US/10767, 701

CURRENT FILING DATE: 2004-01-29

NUMBER OF SEQ ID NOS: 63128

SEQ ID NO 9031

LENGTH: 965

TYPE: DNA

ORGANISM: Sorghum bicolor

FEATURES:

OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS73247_1

US-10-767-701-9031

Alignment Scores:

Pred. No.: 2,29e-52 Length: 965

Score: 634.00 Matches: 143

Percent Similarity: 60.61% Conservative: 17

Best Local Similarity: 54.17% Mismatches: 62

Query Match: 38.63% Indels: 42

DB: 19 Gaps: 10

US-10-630-636-7 (1-306) x US-10-767-701-9031 (1-965)

QY 63 GluGluValArgArgHisTyrGluAlaLeuValGluAspValAlaAlaIleAspAlaGly 82

DB 960 GAGAGGTGACGCGCGCATCGAGCTGCTGTGAGAGACCTTCACGCGCATCGAGTGGCG 901

QY 83 ArgValProLeuProArgTyrAlaGlyGluGluSerAlaAlaProProAspGlyAlaGly 102

DB 900 CGCGTCCGCTCCGAGTACGCGGCTGACGCGGCCCGCAGAGAGGGGTGGCGCGCG 841

QY 103 AlaAlaAlaAlaAlaSerLysAspGlyGlyHisArgArgAspGluArgLysGlyGly 122

DB 840 GGC-----GCGGG-----AAGAGGAGGTGGGGA 814

QY 123 GlyGlyTyrAspGlyGlyLysSerCysSerLys---AlaGluGluGluArgArgLysGly 141

DB 813 GAGAGGACCAACGCGGACAAAGGGGTCCGGAAGTCCGCGAGCAGAGCCGACAGAGAGGC 754

QY 142 IleProTrpThrGluGluGluHisArgLeuPheLeuLeuGlyLysAspLysPheGlyLys 161

DB 753 ATCGCTGACAGAGGACGAGCAGATGTTCTCTCTGAGCTTGAAGAAATGTGGCAA 694

QY 162 GlyAspTrpArgSerLysArgAsnPheValIleSerArgTrpProThrGlnValAla 181

DB 693 GTGACTGGGAGACATCAAGTAATTTGTGATCTCAAGACTCCAACTCAAGTAGCT 634

QY 182 SerHisAlaGlnLysTyrPheIleArgLeuAsnSerMetAsnArgAspArgArgSer 201

DB 633 AGTCATGACAGAAATCTTATCCGCTGATCACTGAACAGAGAGAGCGCGCATCA 574

QY 202 SerIleHisAspIleThrSerValThrAlaGlyAspGlnValAlaAlaGlnGlnGlyAla 221

DB 573 AGTATCATGATCACTCAAGTGTGAACATGAGATGCATCAACGCGACAG-----GCT 520

QY 222 ProIleThrGlyHisGlnAlaThrGlyAsnProAlaAlaAlaAlaLeuGlyProProGly 241

DB 519 CCAATCAAGT---CAACAATGTCAAGCTGA-----AATCTGGG 478

QY 242 MetLysHisHisHisHisHisHisProGlyAlaAlaProProMetProMetTyrSer 261

DB 477 AAACATCTAAGCAATCCCAACAGCCAGCAATTAACCTCCAGCGCTGATAGGC 418

QY 262 AlaAlaProMetGlyHisProValAlaGlyHisMetValProAlaAlaValAlaGlyThrPro 281

DB 417 ACAACA---ATTGACAGCCAGTTGGGCGCTTGTCT---TCAGCTGTGGAATCTCT 364

QY 282 ValValPhePro-----ProGlyHisAlaProTyr----- 291


```

QY 257 MePrOMeCTYrSerAlaAlaProMeGlyHisProValAlaGlyHisMetValProAla 276
Db 967 CCTCTATGTATGGAACACCCGCCATAGTCAAGCA----- 1002
QY 277 AlaValGlyThrProValValPhePro---ProGlyHisAlaProTYrValVal----- 293
Db 1003 GCAGTTGGAACACCACTGAACTCCCACTCCACCTCACTGCTTATGGAGTTCAATGCG 1062
QY 294 ---ProVal---GlyTYrProAlaProProAlaLysMetHis 305
Db 1063 GCTCCAGTCCCTGCTCACTGCTTCTGCTGTCAGCAATGAAC 1104

```

Search completed: July 30, 2005, 01:40:17
 Job time : 729 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 29, 2005, 22:29:33 / Search time 3685 Seconds

(without alignments)
3160.832 Million cell updates/sec

Title: US-10-630-636-7
Perfect score: 1641
Sequence: 1 MTSQAATTTTAAATAAATWTR.....GHAPVVPVGPAPPAKMHQ 306

Scoring table: BLOSUM62

Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2.1/USPRO/seqs/US10630666/runat_27072005_123227_12178/app_query.fasta_1.455
-DB=EST -QFMT=fastp -SUFFIX=est -MINMATCH=0.1 -DOFCL=0 -DOFEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTMT=plco -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10630666 @CGN 1.1 4332 @runat_27072005_123227_12178 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOBJECT -NEG_SCORES=0 -WAIT -DSPELLOCK=100 -LONGLOCK
-DEV TIMEOUT=150 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gse1: *
9: gb_gse2: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	828.5	50.5	884	9	CG257740 OGXCQ48TH
2	816.5	49.8	689	5	BQ483926 WHB3514 C
3	804	48.0	478	7	CF326660 JMT1--06-
C 4	787.5	48.0	759	8	BZ525067 OGART807C
C 5	777.5	47.4	815	6	CA185031 SCBSST309
6	762	46.4	853	9	CG362076 OGICL41TH
7	742	45.2	970	7	CK256709 EST740346
8	742	45.2	973	7	CK254513 EST738150
9	737	44.9	1020	7	CK251993 EST735630

10	717	43.7	952	7	CK294026
11	713.5	43.5	846	7	CO082313
12	694.5	42.3	921	7	CK248989
13	664.5	40.5	936	7	CK292080
14	655.5	39.9	956	7	CK259271
15	655.5	39.9	961	7	CK257717
16	653	39.8	921	9	CG257748
17	635.5	38.7	874	9	CG219919
18	633.5	38.6	1641	3	CNS0A16S
19	633.5	38.5	864	7	CO089242
20	627.5	38.2	1554	3	CNS0A0SA
21	624	38.0	769	7	CK258863
C 22	623.5	38.0	689	9	CG710311
C 23	623.5	38.0	781	7	CV500657
C 24	623.5	38.0	804	9	CG366978
C 25	623.5	38.0	933	9	CG219911
26	611.5	37.3	740	9	CG751932
27	609.5	37.1	1672	3	CNS0A13X
28	607	37.0	974	7	CK290192
29	606.5	37.0	653	9	CG752461
30	606	36.9	628	2	BR356185
31	606	36.9	705	6	CA243768
32	603	36.7	831	7	CO121658
33	597	36.4	621	6	CB003696
C 34	596	36.3	800	7	CK257718
C 35	593.5	36.2	695	7	CK259272
C 36	590	36.0	711	5	BU025086
37	584.5	35.6	599	2	BR433827
38	582.5	35.5	643	6	CA233892
39	578	35.2	778	7	CO081265
40	576	35.1	815	7	CF450963
41	572	34.9	841	7	CO119900
C 42	571.5	34.8	633	5	BO864479
C 43	569.5	34.7	887	7	CO117508
44	560	34.1	678	6	CB001319
45	560	34.1	678	6	CB002201

ALIGNMENTS

RESULT 1
LOCUS CG257740/C
DEFINITION OGXCQ48TH ZM 0.7 1.5 KB zea mays genomic clone ZMBM065G23,
ACCESSION CG257740
VERSION CG257740.1 GI:34164524
KEYWORDS GSS.
SOURCE zea mays
ORGANISM zea mays

REFERENCE
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uetzerback,T.,
1 (bases 1 to 884)
clade; Panicoideae; Andropogoneae; Zea.
Citeck,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Unpublished (2002)
JOURNAL Consortium for Maize Genomics
COMMENT Other GSSs: OGXCQ48TV
Contact: Cathy Whitelaw

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source
1..884
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"


```
Db 180 TCGACGAGGAGCTCCGGGACATCAAGCCCTGTGTGAGAGAGCTGGCGGATTCAC 239
Qy 81 AAGAGTATGValProLeuProArgTyrAlaGluGluSerAlaAlaProProArg 100
Db 240 GCCGCGCGCTCCCTCTCCCGGCTACCGCGGAGAGACTCTCCGCC----- 287
Qy 101 AAGAGTATGValAlaAlaAlaSerLysArgGlyAlaGluGluSerAlaAlaGluGlu 120
Db 288 GCGGCGCGGCTCCCGGCTCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 338
Qy 121 GAGTATGValGlyTyrArgGlyGlyLysSerLysSerLysAlaGluGluGluArgLys 140
Db 339 TTCCAGCTCC-----GATCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 392
Qy 141 GAGTATGValGlyTyrArgGlyGlyLysSerLysSerLysAlaGluGluGluArgLys 160
Db 393 GCGGCGCGGCTCCCGGCTCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 452
Qy 161 LysGlyArgTyrArgSerLysSerLysSerLysAlaGluGluGluArgLys 180
Db 453 AAGGCGAGCTCCGAGAGATCTCCGCAACTCTCCGAGAGAGAGAGAGAGAGAGAGAG 512
Qy 181 AAGTATGValAlaGluGlyTyrArgGlyGlyLysSerLysSerLysAlaGluGluArg 200
Db 513 GCGGCGCGGCTCCCGGCTCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 572
Qy 201 SerSerLysAlaAlaGluGlyTyrArgGlyGlyLysSerLysSerLysAlaGluGluArg 220
Db 573 TTCCAGCTCCAGAGATCTCCGCAACTCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAG 620
Qy 221 AAGTATGValAlaGluGlyTyrArgGlyGlyLysSerLysSerLysAlaGluGluArg 239
Db 621 CCGGCGCGGCTCCCGGCTCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 662
Qy 240 ProGly-----MethylHisHisHisHisHisHisHisHisHisHisHisHis 251
Db 663 CCGGCGCGGCTCCCGGCTCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 689

RESULT 3
CP326660
LOCUS JMT1--06-K20.g1 A3JMT-overexpressing transgenic rice lambda phage 478 bp mRNA linear EST 18-AUG-2003
DEFINITION cDNA library (JMT1) Oryza sativa (japonica cultivar-group) cDNA
clone JMT1--06-K20, Oryza sativa sequence.
ACCESSION CP326660.1 GI:33801577
VERSION 1
KEYWORDS EST.
ORGANISM Oryza sativa (japonica cultivar-group)
SOURCE Oryza sativa (japonica cultivar-group)
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehretaceae; Oryzaeae; Oryza.
REFERENCE 1 (bases 1 to 478)
Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
Large-scale Sequencing Analysis of Rice ESTs
Unpublished (2003)
Contact: Nahm B.H.
Genomics and Genetics Institute, GreenGene Biotech Inc., Division
of Bioscience and Bioinformatics, Yonsei University
Yongin, Kyonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bhnahm@bio.com, bhnahm@bio.myongji.ac.kr.
FEATURES
Source
1..478
Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:39847"
/clone="JMT1--06-K20"
/tissue_type="leaf"
/dev_stage="14 days after germination"
```

```
Db 180 TCGACGAGGAGCTCCGGGACATCAAGCCCTGTGTGAGAGAGCTGGCGGATTCAC 239
Qy 81 AAGAGTATGValProLeuProArgTyrAlaGluGluSerAlaAlaProProArg 100
Db 240 GCCGCGCGCTCCCTCTCCCGGCTACCGCGGAGAGACTCTCCGCC----- 287
Qy 101 AAGAGTATGValAlaAlaAlaSerLysArgGlyAlaGluGluSerAlaAlaGluGlu 120
Db 288 GCGGCGCGGCTCCCGGCTCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 338
Qy 121 GAGTATGValGlyTyrArgGlyGlyLysSerLysSerLysAlaGluGluGluArgLys 140
Db 339 TTCCAGCTCC-----GATCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 392
Qy 141 GAGTATGValGlyTyrArgGlyGlyLysSerLysSerLysAlaGluGluGluArgLys 160
Db 393 GCGGCGCGGCTCCCGGCTCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 452
Qy 161 LysGlyArgTyrArgSerLysSerLysSerLysAlaGluGluGluArgLys 180
Db 453 AAGGCGAGCTCCGAGAGATCTCCGCAACTCTCCGAGAGAGAGAGAGAGAGAGAGAG 512
Qy 181 AAGTATGValAlaGluGlyTyrArgGlyGlyLysSerLysSerLysAlaGluGluArg 200
Db 513 GCGGCGCGGCTCCCGGCTCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 572
Qy 201 SerSerLysAlaAlaGluGlyTyrArgGlyGlyLysSerLysSerLysAlaGluGluArg 220
Db 573 TTCCAGCTCCAGAGATCTCCGCAACTCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAG 620
Qy 221 AAGTATGValAlaGluGlyTyrArgGlyGlyLysSerLysSerLysAlaGluGluArg 239
Db 621 CCGGCGCGGCTCCCGGCTCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 662
Qy 240 ProGly-----MethylHisHisHisHisHisHisHisHisHisHisHisHis 251
Db 663 CCGGCGCGGCTCCCGGCTCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 689

RESULT 4
B2525067/c
LOCUS B2525067 759 bp DNA linear GSS 16-DEC-2002
DEFINITION OGAH807C ZM2 0.7.1.5 KB Zea mays genomic clone ZMMBMA0064M16,
genomic survey sequence.
ACCESSION B2525067
VERSION B2525067.1 GI:27062615
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 759)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uetzelback,T.,
Reanick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nuneberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genome
Unpublished (2002)
TITLE JOURNAL
COMMENT Contact: Cathy Whitelaw
TIGR
```



```

Db          | 666 CTATCTCCTGCGGTGGATTACCCGCCGCACCAGCAAGAATGCACCCAG   714
|||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
RESULT 7    | CK256709                               970 bp      mRNA       linear     EST 30-JUL-2004
CK256709    | ESI740346 potato callus cDNA library, normalized and full-length
DEFINITION  | Solanum tuberosum cDNA clone POC46_5' end, mRNA sequence.
ACCESSION   | CK256709
VERSION     | CK256709.1 GI:39813689
SOURCE      | EST.
ORGANISM    | Solanum tuberosum (potato)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; Lamiales; Solanales; Solanaceae; Solanum.
REFERENCE   | Buell,C.R., Hart,A., Zisman,V., Karmychev,S.A. and Baker,B.
AUTHORS     | Generation of ESTs from potato callus tissue
TITLE        | Unpublished (2003)
JOURNAL      | Contact: Robin Buell
COMMENT      | The Institute for Genomic Research
              | 9712 Medical Center Dr, Rockville, MD 20850, USA
              | Email: potcdo-array@tigr.org
              | Clones can be requested from the University of Arizona Genomics
              | Institute via http://genome.arizona.edu/orders/.
              | Seq primer: ATT TAG GTG ACA CTA TTAG.
              | Location/Qualifiers
FEATURES             |
source                | 1..970
                     | /organism="Solanum tuberosum"
                     | /mol_type="mRNA"
                     | /cultivar="Kennebec"
                     | /db_xref="taxon:4113"
                     | /clone="POCCV46"
                     | /tissue_type="callus"
                     | /lab_host="DHIOB-Tona"
                     | /clone_lib="potato callus cDNA library, normalized and
                     | full_length"
                     | /note="vector: PCWVSporte.1; Site.1: EcoRI; Site.2: NotI;
                     | supplier: RNA was isolated from Solanum tuberosum var.
                     | Kennebec callus tissue grown on solid media."
ORIGIN
Alignment Scores:
Pred. No.:           1,14e-46               Length:         970
Score:               742.00                 Matches:         169
Best Local Similarity: 63.24%               Conservative:    34
Overall Similarity:   52.65%               Mismatches:     62
Query Match:         45.22%               Indels:         56
DB:                  7                      Gaps:           16

US-10-630-636-7 (1-306) x CK256709 (1-970)
OY      | 13 AAlAAlAAAlAAATrPThArRGGuBaSPAPlySAIAPhEGluASnALAEUAlAA 32
       | ::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      | 61 TCATCTTCAAGTAgTGTTGGAGCAAGAGGAAGAAAAGCTTTTgAGATCCATGTC TG 120
OY      | 33 CySaLaAlaProProProLaAspgLygLYalAProASPaspApTrP----- 48
       | ::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      | 121 -----CACTCGGGTGAAGACTTC 138
OY      | 49 -----PheAlAlAlAEuAlAlaserValPrOGlyAlArSGerAlagLUglUval 65
       | :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      | 139 GAACAACAATGGAGACAGCTTGTCTTARGGTCCCC--ACAAGAACCATTATGAGCTT 195
OY      | 66 ARGARSHStETrgLUALAEUVAlGUaSPVAAlAlAlAlAsPALaglYarGalPro 85
       | ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      | 196 AAAGAACATTACCAAGCTACTTTTGGAGAGTGTGGAGCCATTGGAAGCCAAGCTTCA 255
OY      | 86 LeUPrArgTrYalAGlYGLUGLu---SerAlAlaPrOProASPglYAlAGlYlaAla 104
       | ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      | 256 ATACCAATTACAAGAGAGAAAGAGCTTCTTCCTCACTTAAGAAAGATTAACTAGGA 315

```

QY		105	AlaAalaaISerLYAspGlyGIYHisArgIrgApRguIA-gLyGLYglYglYglY	124
Dd		316	TATCTGGGTCAAGTGAACCGA-----CGGCCTCCAAATTGCGTTATTCAAATGA	366
QY		125	TyrAsp-----gLygLySserCySerIysAlaglu	135
Dd		367	TTTTCAGTAGACACCATTGCCATTCGGCATGGTGAAAAAGGAAATTCTAGCTGAA	426
QY		136	GlNGluA-gA-gLyLleProTrPThrGUlguLuhIeArgleurPheLeuleuGLY	155
Dd		427	CAGAAGAACAAAAGCATACCGGACTAAGAAAGAACATTAAGTTTCTTTTGGTAGGT	486
QY		156	LeuAspLYSPhegLyLYeGLYASptIrgAsrSerIIeseraGSanPhenValIIeserArg	175
Dd		487	TTAGACAATTTGGAAAGAGATTGGAGAAATATTAAAGAAATTTGTTATCTATCTGA	546
QY		176	ThrProthrGlnValaIsaSerHislaInlunYSrYrPheIIeargleuAsnSerMerAn	195
Dd		547	ACACCAACACAAGTAGTACTAGTCAGCCAGAAATCTTATTCTGTTGAATTCATGAAT	606
QY		196	ArgAspA-gA-gA-gSerSerIIeHIsApRIleThrservalThraGLYAoRglNvaI	215
Dd		607	AGAGATGAAGAAGGTCAAGTATTCACACATTCACAACTATCAATATATGAGAGACGTT	666
QY		216	AlaAlaIngInglYAlaProIlleThrGlyHisgInlalathrNgLYAsnProLa---Ala	234
Dd		667	TCAACTCATCAA--GCTTCCAATTACAGGCCAACAGGTGAATCCGAAATCCATCAATCA	723
QY		235	AlaAlaleugLYProProGlyMetLYHisHis-----HisHIsHIsHIsProGLY	251
Dd		724	GGTGCTTTGA---CCCTCAGTCAAGCACAGAACCAACCAATATGCATAGTAGGGC	780
QY		252	GLyAlaProProPrometProMetLYSeralAlaAPrometGLYHisProValAla---	270
Dd		781	-----ATGATAC---GATGCTCCAGTAGGGTCAACCCCTCCCTGCA	816
QY		271	-----GLYHisMeLVaIProAlalaValaGlyThrProvalIvalPheProProGLY---	287
Dd		817	CCGCCAAGCCACTTGCC--TCAGCAATCGGAACACCGGTTATTCCTTCATGATGAC	873
QY		288	HIsAlaProTYrValaIvalProValIGLYr-----ProAlaProProAlaIysMethHis	305
Dd		874	CATCTCCTTATGTACTTCACATTGCCAATCTTAACGCCACCGCCACACCAACCATTGAC	933
QY		306	Gln 306	
Dd		934	CAG 936	
<hr/>				
RESULT 8				
CK254513				
LOCUS		973 bp	mRNA	linear EST 30-JUL-2004
DEFINITION		EST738150 potato callus cDNA library, normalized and full-length		
VERSION		Solanum tuberosum cDNA clone FOCDD92 5' end, mRNA sequence.		
ACCESSION		CK254513		
KEYWORDS		CK254513.1 GI:39810497		
SOURCE		EST		
ORGANISM		Solanum tuberosum (potato)		
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
AUTHORS		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicts;		
TITLE		Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.		
JOURNAL		Generation of ESTs from potato callus tissue		
COMMENT		Unpublished (2003)		
		Other ESTs: EST738151		
		Contact: Robin Buell		
		The Institute for Genomic Research		
		9712 Medical Center Dr, Rockville, MD 20850, USA		
		Email: potcoto-array@tigr.org		
		Clones can be requested from the University of Arizona Genomics		
		Institute via http://genome.arizona.edu/orders/ .		
		Seq primer: ATT TAG GTG ACA CTA TAG.		


```

Oy 123 -----GlyGlyTyr-----AspGlyGly 128
Db 366 TCTTTTGTAGTATGTAATGATTTTCAGCTAGTATCAATGAATTCGCCAAGAGGA 425
Oy 129 LysSerCysSerLysAlaGluGlnLysArgArgLysGlyLeuProThrGluGlu 148
Db 426 AAGGAAAGTTCAAATTCGAAACGAAAGAAAGAAAGAAATCTTGACCTCAAGAGG 485
Oy 149 HisAlaGluPheLeuLeuGlyLeuAspLysPheGlyLysGlyAspTrpArgSerIleSer 168
Db 486 CATAGCTTTCTCTGCTGTTTATGATGATGCGAAGGAGGAGCTGGAGAAAGTATTTCA 545
Oy 169 ArgAsnPheValIleSerArgThrProThrGlnValAlaSerHisAlaGlnLysTrpPhe 188
Db 546 AGGAATTTATGATTTCTAGAACACCACTCAAGGTGCTAGTACAGCCCAAGTACTT 605
Oy 189 IleArgLeuAsnSerMetAsnArgAspArgArgArgSerIleHisAspIleThrSer 208
Db 606 ATCCGATTAAATTCATGATGATGAGATGAGAGAGTCCAGCATGACATTCATCAAGT 665
Oy 209 ValThrAlaGlyAspGlnValAlaGlnGlnGlyAlaProIleThrGlyHisGlnAla 228
Db 666 GTCAACAAATGCG--GAGGTGCTCAAAATCAG--ACTCTTATAACAGGCAAAACACC 719
Oy 229 ThrGlyAsnProAlaAlaAlaLeuGlyProProGlyMetLysHisHisHisHis 248
Db 720 AATACAGCCCATCAATGACAGTACACCACTGTCGGGCAAAAGAACAGCCCAATATG 779
Oy 249 HisProGlyGlyAlaProProPheMetProMetLysSerAlaAlaProMetGlyHisPro 268
Db 780 CAT-----GGTTGAGTATGAT--GGTCTCCATGAGTCTCATCG 818
Oy 269 ValAla-----GlyHisMetValProAlaAlaValGlyThrProValAlaPhePro 285
Db 819 GTCCGTGCTCCACCGGGGACACCGCTG--TCAGCAGTCGGAACACCGGTCTCTCT 875
Oy 286 ProGly---HisAlaProThrValAlaProValGlyTyrProAlaProPro 301
Db 876 CTTGACATCATCCACCTTATGTTCTTCCAGCTGATACACATGCGCGG 926

```

```

RESULT 11          846 bp  mRNA      linear  EST 15-JUN-2004
LOCUS             CO082313
DEFINITION GR_Ea46N19.r GR_Ea Gossypium raimondii cDNA clone GR_Ea46N19.3',
mRNA sequence.
ACCESSION         CO082313
VERSION           CO082313.1 GI:48751794
KEYWORDS          EST.
SOURCE            Gossypium raimondii;
ORGANISM          Gossypium raimondii;
                  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                  rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE         1 (bases 1 to 846)
AUTHORS           Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
                  Udell,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
                  Wing,R.A.
TITLE             Global assembly of Cotton ESTs
JOURNAL           Unpublished (2004)
COMMENT           Contact: Rod A. Wing
                  Arizona Genomics Institute
                  The University of Arizona
                  Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
                  Tel: 520 626 9595
                  Fax: 520 621 1259
                  Email: http://genome.arizona.edu
                  Plate: 46 row: N column: 19.
FEATURES          Location/Qualifiers
source            1..846
                  /organism="Gossypium raimondii"
                  /mol_type="mRNA"
                  /db_xref="taxon:29730"

```

```

/clone="GR_Ea46N19"
/cisue_type="whole seedlings"
/dev stage="first true leaves"
/lab_host="DH10B"
/clone_id="GR_Ea"
/notes="vector: pCMVSPORT-6.1; Site 1: NotI; Site 2:
EcoRV; library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGT. More glycerol clones held in -80."

ORIGIN
Alignment Scores:
Pred. No.: 1,44e-44 Length: 846
Score: 713.50 Matches: 163
Percent Similarity: 63.82% Conservative: 31
Best Local Similarity: 53.62% Mismatches: 61
Query Match: 43.48% Indels: 49
DB: 7 Gaps: 14

US-10-630-636-7 (1-306) x CO082313 (1-846)
Oy 13 AlAlaAlaAlaAlaTTPThArgGluAspLysAlaPheGluAsnAlaLeuAla--- 31
Db 2 GCAGGAAGTGTAGTTGAGCAGGAGGACATGAGAGGCAATTGAAACGCAATACCAATG 61
Oy 32 -----AlaCysAlaAlaProProAlaAspGlyAlaProAsp 45
Db 62 CATTGATAGAGAAAGAGAAATG-----GAA 88
Oy 46 AspAspTrpPheAlaAlaLeuAlaAlaSerValProGlyAlaArgSerAlaGluGluVal 65
Db 89 GAGCAATG---GAAAGATTGCTTCAACGGTCT--ACTAAAGCTTGGAGAAATG 142
Oy 66 ArgArgHisTyrGluAlaLeuValGluAspValAlaAlaIleAspAlaGlyArgValPro 85
Db 143 AAGCTTCACTAGCAATTTATGTCGAGATGATCTGCAATAGAGGCTGTCATGTTCCA 202
Oy 86 LeuProAspTyrAlaGlyGluGlu-----SerAlaAlaProProAspGlyAlaGlyAla 103
Db 203 CTTCTCTGTATTAAGGAGAAAGAACCTTCTTCTGCTCAAGATTAATTTTCATGACCT 262
Oy 104 AlAlaAlaAlaSerLysAspGlyGlyHisArgArgAspGluArgLysGlyGlyGly 123
Db 263 TCCATGCTCTTAATTCGAT-----AGAAATCCAATTCGGCTTACGGAAT 310
Oy 124 GlyTyrAsp-----GlyGlyLysSerCysSerLysAla 134
Db 311 AGTTTTCGGGATTAACTTGAATTCAGTACGACATGAGGAGAAACAGAGTTCCGAGTCC 370
Oy 135 GluGlnGluArgArgLysGlyLysProThrProThrGluGluGluHisArgLeuPheLeu 154
Db 371 GACCAAGAAAGAAAGAAAGAAATCCCATGACTGAAGAAAGCAATAGGATTTCTTACTT 430
Oy 155 GlyLeuAspLysPheGlyLysGlyAspTrpArgSerIleSerArgAspPheValIleSer 174
Db 431 GCTTTCAGCACTTTGGAAAGAGATTCGAGAACATTTCAAGAACCTTTGTGATATCG 490
Oy 175 ArgThrProThrGlnValAlaSerHisAlaGlnLysTyrPheIleArgLeuAsnSerMet 194
Db 491 AGAACTCCGACGACAGGTGCTAGCATGACAAAGATTTTATATACGCTTAATTCGATG 550
Oy 195 AsnArgAspArgArgArgSerSerIleHisAspIleThrSerValThrAlaGlyAspGln 214
Db 551 AATAGAGACCGCGCGGTCCAGCATCCACGACATCAGAGGTGTTAACATATGGGAT--- 607
Oy 215 ValAlaAlaGlnGlnGlyAlaProIleThrGlyHisGlnAla---ThrGlyAsnProAla 233
Db 608 -----ACGTCTCATCAAGCCCTATTCACCGCCCAACGATTAACAACTCACCCGGT 661
Oy 234 AlAlaAlaLeuGlyProProGlyMetLysHisHisHisHisHisHisProGlyGlyAla 253
Db 662 GCAGCGGTCAATGGG---CAATCACTAGTAAGCAGAGGCTCAGCGCAT----- 706

```

[illegible]

Accession	Version	Source	Organism	Reference	Title	Journal	Comment
Db	183	GAAACGACAAATGGAGCAGATTGTTCTTATAGTGTCCCC---ACAAAGAACCATGTATGAGCTT	239				
Qy	66	ArgArgHisIstYrGluAlaLeuValGluAspValAlaAlaIleAspAlaGlyArgValPro	85				
Db	240	ANGGAAACATTTACACAGCTTACTTTTGGAAAGATGTTCCAGCATTTGAAGCAGGCCAACTTCCA	299				
Qy	86	LeuProArgTyrAlaGlyGluGlu---SerAlaIaProProAspGlyAlaGlyAlaIa	104				
Db	300	ATTACCAATTTACAAAGAGAGAAAGAGCTTCTTCTTCACTTAAGAAAGATTAACTTAGA	359				
Qy	105	AlaAlaIaIaSerIyAspGlyGlyIyHisIaArgIaAspGluArgIyGlyGlyGlyGly	124				
Db	360	TATCTGGGTCAGTGGACGGA-----CGGGGCTCCAAATTGNGTTATTCAAATGGA	410				
Qy	125	TyrAsp-----GlyGlyIySerCysSerIyAlaGlu	135				
Db	411	TTTTTCAGTACGACCCATGACCCGATTTGGCATGGTGGAAAAAGGAAATTCAGGCTGGA	470				
Qy	136	GlnGluArgArgGlyGlyIleProTyrThrGlnGluGluGluIyHisArgIyLeuLeuGly	155				
Db	471	CAAGAGAGACGAAAGGATATCCGTGGCTAAGAGAAACATAGGTGTGTTTGGCTAGCT	530				
Qy	156	LeuAspIyAspPheGlyIyGlyIyAspTyrArgSerIleSerArgAspPheValIleSerArg	175				
Db	531	TTTACACAAATTTGGGAAAGAGATTTGGAGAGATTTTCAAGGAATTTGTTATATCTCGA	590				
Qy	176	ThrProThrGlnValAlaIaSerHisAlaGlnIyIyTyrPheIleArgLeuAspSerMetAsn	195				
Db	591	ACACCAACACACAAAGTAGTACATGACGCCAGAAATACCTTATTCGTTGAATCCCATGAT	650				
Qy	196	ArgAspArgArgArgSerSerIleHisAspIleThrSerValThrAlaGlyAspGlnVal	215				
Db	651	AGAAATGAGAAAGAGTCAAGTATTCAGACATTCACAACTATCAATATATGAGAGACGTT	710				
Qy	216	AlaAlaGlnGlnGlnGlyAlaProIleThrGlyHisGlnAlaThrGlyAspProAla---Ala	234				
Db	711	TCACATCATCA---GCTCCAAATTTACAGCCCAACAGTGATCCGAATCCATCAATATCCA	767				
Qy	235	AlaAlaIaLeuGlyProProGlyMetIyHisHis-----HisHisHisIleProGly	251				
Db	768	GCTGCTCTTGGA---CCCTCAGTCAAGACAGAACCCCAACCAATATCATATGATAGTGGGC	824				
Qy	252	GlyIaIaProProPromePromeTyrSerMetIySerAlaIaProMetGlyHisProValAla---	270				
Db	825	-----ATGTAC---GATGCTCGATGGGTGCACCCGTCCTGCA	860				
Qy	271	-----GlyHisMetValProAlaAlaValGlyThrProValAlaPheProProGlyHis	288				
Db	861	CCGGCAAGCCACCTTGGC---TCAGCAGTCCGAAACCGGTTATGCTTCCCTCATATGAGAC	917				
RESULT 13							
LOCUS	CK292080	938 bp	mrna	linear	EST 02-AUG-2000		
DEFINITION	EST754794	Nicotiana benthamiana mixed tissue cDNA library,					
		normalized, full-length	Nicotiana benthamiana	cDNA clone	NBMC676	5	
		end, mRNA sequence.					
ACCESSION	CK292080						
VERSION	CK292080.1	GI:39873168					
KEYWORDS	EST.						
SOURCE	Nicotiana benthamiana						
ORGANISM	Nicotiana benthamiana						
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;						
	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;						
	asterids; lamiales; Solanales; Solanaceae; Nicotiana.						
REFERENCE	Buell,C.R., Hart,A., Zismann,V., Kararycheva,S.A., Day,B.,						
AUTHORS	Staskiewicz,B., Jin,H. and Baker,B.						
	Generation of EST sequences from Nicotiana benthamiana						
	Unpublished (2003)						
	Other ESTs: EST754795						
	Contact: Robin Buell						
	The Institute for Genomic Research						
	9712 Medical Center Dr, Rockville, MD 20850, USA						


```
Oy 156 LeuAplysPheGlyLeGlyAspTrpArgSerIleSerArgAsnPheValIleSerArg 175
Db 623 TTAGCAAAATTGGGAAAGAGATTGAGAAATTTCAGGAATTTGTATATCTCGA 682
Oy 176 ThrProThrGlnValAlaSerHisAlaGlnIleTyrPheIleArgLeuAsnSerMetAsn 195
Db 683 ACACCAACACAGTAGTACTAGTCATGCCAGAAATACCTTATTCGTTGAATTCATGAAT 742
Oy 196 ArgAspArgArgSerSerIleHisAspIleThrSerValThrAlaGlyAspGlnVal 215
Db 743 AGAGATAGAGAGAGCTCAAGTATTCACGACATTCACAGTATCAATATGAGAGACGTT 802
Oy 216 AlaAlaGlnGlnGlyAlaProIleThrGlyHisGlnAlaThrGlyAsnProAla---Ala 234
Db 803 TCAACTCATCAA---GCTCCAATTACAGGCCAACAGTGAAATCCGAATCCATCAATCCA 859
Oy 235 AlaAlaLeuGlyProProGlyMetLeuHisHis-----HisHisHisHisProGly 251
Db 860 GCTGCTCTTGA---CCCTCAGTCAGACAGAACCCAAACCCATATGATATGATGAGC 916
Oy 252 GlyAlaProProMetProMetTyrSerAlaAlaProMetGlyHisProValAla 270
Db 917 -----ATGTAC---GGTGCTCCGATGGGTCAACCCCGCTCGCT 949
```

Search completed: July 30, 2005, 01:24:41
Job time : 3698 secs

This Page Blank (uspto)

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☒ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☒ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

This Page Blank (uspto)